Annual Review of Political Science 2012 XX 1056-8700/97/0610-00

Richardson in the Information Age: GIS and Spatial Data in International Studies

Kristian Skrede Gleditsch

University of Essex and Peace Research Institute Oslo, ksg@essex.ac.uk

Nils B. Weidmann

Peace Research Institute Oslo, nils.weidmann@gmail.com

Key Words Geographic Information Systems, Spatial Data, Spatial Disaggregation, Spatial Processes, Spatial Diffusion

Abstract There is an enduring interest in how geographical features influence political interactions and outcomes, and many key factors highlighted in international relations and cross-national research vary spatially within countries. Starting with the pioneering research of Richardson, we show how Geographic Information Systems (GIS) technology and the increasing availability of spatial data can provide new opportunities to answer old and new questions. We focus on key motivations for using spatial disaggregated data and show how such data can help advance core research questions, drawing on examples from the study of violent conflict.

CONTENTS

Introduction

> [G]eography and the character of the ground bear a close and everpresent relation to warfare, . . . both as to its course and to its planning and exploitation (von Clausewitz 1832/1984:348)

Geographical concepts such as geopolitics, or the notion that territorial and geographical features shape strategic considerations and decisions, have been central to international studies from the outset. Early theoretical contributions emphasized both "chaps", in the sense of the key decisionmakers and strategic interaction, as well as "maps", in the sense of how the geographical location and context of international politics influence their aims and actions. The pioneer of modern conflict research, Lewis Fry Richardson, made many well-known important contributions to modelling interactions such as arms races using differential equations (Hess 1995; Nicholson 1999; Richardson 1960*a*). Less well known is his work on geography, where Richardson (1960*b*; 1992) also made important contributions to topics such as the relationship between borders and conflict, measures of territorial properties such as "compactness" and their relationship to conflict and defense, as well as a number of interesting observations on the political implications and origins of borders. For example, Richardson noted how administratively determined internal borders tended to look very different from "natural" external borders. While the latter tend to follow physical features such as rivers or mountain ranges, the former often take the form of straight lines clearly drawn directly on a map, usually without regards for natural features. For example, Richardson pointed out that there were no instances of four independent states meeting in a single point, such as the Four Corners area of the

United States. He attributed this to the role of warfare in shaping borders and the difficulty of maintaining border arrangements that would be difficult to defend militarily. The Caprivi strip, a protruding part of northeastern Namibia, is sometimes held up as a contemporary counterexample to Richardson's observation, although Namibia and Zimbabwe do not appear to be contiguous although both border the Zambezi river. However, consistent with Richardson's core intuition, this strangely shaped area, emerging from complex treaties between the UK and Germany, has seen considerable conflict and contention.

Richardson had very limited tools at his disposal when writing in the 1930s, and most of his geographical computations were done by hand. We have seen large advances in mathematical modelling and analysis of data in international studies over the latest decades (Morrow 2000; Ward 2010), in part following advances in modern computing and the increasing popularity of formal modeling in other fields. However, social scientists have been much slower in picking up on advances in Geographic Information Systems (GIS) and the increasing availability of disaggregated, high resolution spatial data. In this article, we review existing research that has picked up the gavel from Richardson using the tools of the information age, and some of the possibilities that this line of research opens up for advancing international studies and cross-comparative research.

We focus on three important motivations for using spatial data, namely (i) using spatially disaggregated information to generate new data that would otherwise not be available, (ii) spatial disaggregation, where the spatial variation within units is a key focus of interest, and (iii) cases where there is a substantive focus on spatial processes and dynamics. As we cover an emerging research area and concepts that may be unfamiliar to many readers, our discussion will provide more illustrative examples on applications to substantitive questions and avenues for future research than a more conventional review. Although the examples reflect our own interests in violent conflict, they help illustrate more general points relevant to a wide range of research questions. Still, our overview must necessarily be selective. Since we focus on empirical data and applications, we will not attempt to review the range of related conceptual or non-quantitative research in human geography and political geography (Ethington and McDaniel 2007; Cox, Low and Robinson 2007; Flint and Taylor 2007). Given our specific focus on GIS and spatial data, we cannot offer full justice to the wealth of research that has examined issues related to Richardson's original questions using other methods, including important work on borders and conflict (Siverson and Starr 1991) and the political and economic implications of "artificial borders" (Alesina, Easterly and Matuszeski 2011). We will first review properties of spatial data and some important data sources relevant to central issues and questions in international studies and comparative politics. We will then proceed to more detailed examples of how spatial data can provide important new insights, both in terms of how spatial disaggregation can enhance theory-measure correspondence and illuminate important spatial variation, as well as in terms of modelling important spatial processes.

Core Concepts of Geographic Information Systems

The broad term "Geographic Information System" is used to describe a family of software tools that allow for the collection, visualization and analysis of spatial data. In this paragraph, we first provide a brief introduction to the nature of spatial data, and later turn to the analysis of this type of data using GIS.

Spatial data stored and processed in a GIS are typically represented following one of two major approaches. The first one is the *vector data* approach. A vector dataset includes real-world geographic features as separate entities. For example, cities can conveniently be represented as (x, y) dots (see Figure 1, left), while more complex geometries such as rivers or lakes require sets of points connected as a line (Figure 1, center) or a polygon (Figure 1, right). Our `cshapes` dataset (Weidmann, Kuse and Gleditsch 2010), which contains time-varying country borders, is another example of a vector dataset. In `cshapes`, each country corresponds to a polygon that shows the borders of this country at a given point in time.

Along with the spatial features contained in a vector dataset, non-spatial information can be attached through a separate *attribute table*, where each row is linked to one feature in the dataset. For example, in our `cshapes` dataset, the attribute table contains the name of the country as well as country identifiers according to different coding schemes. Also, the attribute table allows us to specify lifespans of individual polygons. This way, a country polygon is succeeded by another if the country borders change because of secession or territorial conquest.

A second type of data representation in GIS is the raster data approach. A GIS raster is a representation of a continuous spatial variable. The entire space is divided into equal-sized cells, and the value of the variable of interest is measured for each of these cells. Environmental variables are typically represented in raster formats. Figure 2 shows an example of how territorial elevation can be stored in a raster dataset. Each of the cells in the raster (right) contains a single elevation value, where darker shading correspond to cells of higher altitude. Collier and Hoeffler (2004) and Fearon and Laitin (2003), for example, use an index of the proportion of a state that is composed of “mountainous terrain” based on raster data on elevation. Also, spatial datasets derived from satellite images (see below) are often distributed in a raster format, since the satellite's measurement is continuous over space.

Both raster and vector data require us to specify how information from the three-dimensional surface of the earth should be projected to a two-dimensional map surface. Since it is impossible to represent a sphere on a planar surface, all maps must inevitably introduce some distortion, and researchers must chose which properties of the sphere they want to preserve (for example, area, distance, shape, or direction) as no map can preserve all these properties at the same time. A full discussion of map projections and coordinate systems is beyond the scope of this article, and we refer the reader to the standard geographical literature on this topic (Longley et al. 2010).

The choice of the right data model (vector or raster) depends on a number of factors. First, vector and raster datasets differ as regards the precision by which they can represent spatial information. In a raster dataset, precision is limited by the size of the raster cells (the resolution), and it is not possible to move below this resolution. In a vector dataset, however, individual features can be stored using an arbitrary number of points with coordinates of high precision, such that the dataset's precision is—theoretically—unconstrained. However, as first noted by Richardson (1992), some geographical features such as coastlines have fractal properties that are difficult to measure or describe and dependent on the resolution (this is now known as the "Richardson effect", see Mandelbrot 1967). Second, vector data are more flexible. While a raster represents a single continuous variable over space, a vector dataset can hold more than one variable that applies to its spatial features. For example, it is possible to append a new column to the `cshapes` attribute table, such that each country polygon holds its democracy score in addition to the existing variables. However, this flexibility comes at a price: The third major difference between the two data models is computational costs for processing. Because of their high precision and versatility, computations based on vector data quickly reach their limits. This is why most of the large simulation models based on GIS data (e.g. climate and weather models) are typically run in a raster structure.

The primary features of a GIS are (i) to help collect and modify spatial data, and (ii) to analyze it. For most social scientists, GIS analysis tends to be equated with visualization and software to generate maps. It is true that much can be learned from plotting data on a map and inspecting it. John Snow's map of the 1854 cholera epidemic in London provides a famous historical example of how maps can generate new insights or confirms suspicions. The map indicated that the individual incident clustered in a manner suggesting a single water pump in Broad Street, Soho as the likely source, confirming Snow's initial theories regarding the role of contaminated water and providing evidence against the miasma theory that cholera was caused by "bad air"(Brody et al. 2000). Maps are a powerful way of communicating information, and clearly underused in the social sciences. In some cases, visualization can help uncover details that would not be visible in the tabular data. For example, an attempt to create a movie of annual maps to display the diffusion of democracy as evidenced in the Polity data in 1996 uncovered errors in the recorded values that had previously escaped researchers (Gleditsch and Ward 2005).

However, GIS analysis extends beyond creating maps. In our view, the key promise of GIS analysis for research in international studies lies precisely in the ability to run operations to calculate statistics of interests or reflect spatial variation. Some of the analysis operations work on single dataset (or "layer") as input. For example, computing minimum distances between pairs of countries—as implemented in our cshapes package for R (Weidmann and Gleditsch 2010)—requires only one input layer of country borders. More complex operations (so-called "overlays") make use of more than one layer. These operations take advantage of the spatial co-occurrence of information contained in different datasets. For example, if we know where a country's administrative units are located, and overlay this with information about territorial elevation, we can find out which of these districts should be considered mountainous by some specific criteria. The variety of overlays and other operations that can potentially be conducted in a GIS is huge, which prevents us from a detailed discussion. However, we present a number of examples below that will convey more intuition about the potential of GIS analysis in our field.

Sources of GIS Data

With the rapid growth in the use of GIS over the last decade, there now exist a wealth of GIS datasets available that cover issues relevant to researchers in international studies and comparative politics. Furthermore, it is straightforward to collect new GIS data for the purpose of spatial analysis. In this section, we review both existing datasets and describe ways of collecting new spatial data for a research project.

Existing Spatial Datasets

As we described above, spatial data can be represented either in a vector or a raster format. Vector formats are typically used for discrete spatial entities, while rasters represent a continuous variable over space. In the following, we briefly describe datasets of both formats, distinguishing between four thematic categories: *Political* datasets on national and subnational borders, *demographic* datasets that map the population distribution or of particular population group, *socio-economic* data that provide spatial information about features such as economic activity, the distribution of specific resources, or transportation infrastructure, and our last category includes spatial datasets with information on *environmental* factors. In addition to the (mostly global) datasets described below, there exist a number of national and subnational spatial datasets for specific countries. These datasets are often more detailed and of better quality than their global counterparts, and can thus be extremely useful for studies focusing on a specific country or region (Snyder 2001). However, due to limited space it is impossible for us to provide a comprehensive review of such datasets here.

POLITICAL National and subnational political borders are often available as vector datasets. Even if the precise extent of these entities is not of primary interest, it becomes possible to link other variables of interest—which often are measured at the level of political entities—to their spatial referents, and thus create maps, estimate spatial statistics of the quantities of interest, or analyze spatial processes such as diffusion. Our `cshapes` dataset provides historical country borders as vector polygons for the post-World War II period (Weidmann, Kuse and Gleditsch 2010). Each polygon is labeled with the country identifiers according to the most frequently used reference systems in the discipline (*Correlates* *of War* and the *Gleditsch & Ward* 1999 state list), which makes it possible to merge other country-level variables to the dataset such as for example the Polity democracy measures (Jaggers and Gurr 1995) to produce maps of the distribution of democracy. In addition, the associated `cshapes` R package allows the user to compute derived measures from the country polygons, such as the minimum distance between countries (Weidmann and Gleditsch 2010). Other country border datasets exist, but are mostly incompatible with coding conventions in the discipline (Patterson and Kelso 2011).

Similarly, subnational borders are available as GIS datasets. One of the most comprehensive ones is the *GADM Database of Global Administrative Areas*, available at `http://www.gadm.org`. GADM provides vector datasets of the first- and sometimes second-level administrative units of countries around the world. What limits the applicability of this dataset for cross-national studies is the lack of a temporal dimension; GADM always attempts to provide the latest administrative configuration of a country, and does not show the various changes to this configuration over time. Other sources such as *Statoids* (Law 1999) trace this history and provide maps, but not in a format that can be processed in a GIS. Statoids was used by Cunningham and Weidmann (2010) to create a dataset of first-level administrative units in 1991, which is publicly available in a vector format. A new GIS dataset on federal subunits is being created by Deiwiks (2010), which tracks boundaries of federal units over time in a GIS vector format. Similar to our approach in `cshapes`, the dataset assigns "lifespans" to polygons to capture this evolution.

Research on political violence has generated a host of different geographic datasets, some of which we review here. One of the first attempts to spatially reference civil war was conducted by Buhaug and Gates (2002), who code a "conflict zone" as the smallest circle that surrounds all violent events in a given country and year. The *Conflict Site* dataset (Raleigh et al. 2006) uses this approach and provides information on conflict zones for all conflicts in the (non-spatial) Uppsala/PRIO *Armed Conflict Dataset* (Gleditsch et al. 2002). Since the radial approximation of conflict zones is only a rough approximation of the real area affected by violence in a conflict, Buhaug and Rød (2006) replace the circles by using polygons that can take arbitrary shapes, are clipped to country boundaries and only encompass the country experiencing conflict, and thus can approximate the conflict zone better. The original Buhaug and Rød (2006) data covered only Africa and have not been publicly released. However, a revised version of such polygons have been developed as an extension to the Conflict Site data by Hallberg (2011), and will soon be made publicly available.

By construction, the conflict zone representation is a simplification of the location of fighting in a civil war. Many researchers have attempted to move to a more precise level of analysis and to focus on the individual events that make up "a war" or conflict episode. Currently, attempts are underway to report fighting in these conflicts at the incident level, where each incident is tagged with temporal and spatial coordinates, thus employing a point vector representation as described above. One these conflict event datasets is the *Armed Conflict Location and Event Dataset* ACLED (Raleigh et al. 2010), which makes incident collections for many contemporary conflicts available online (`http://www.acleddata.com`). A similar effort is being conducted at the Uppsala Conflict Data Project (Melander and Sundberg 2011). Other major datasets in the discipline have also been geo-referenced. The MIDLOC dataset (Braithwaite 2010) reports the onset of each episode in the *Militarized Interstate Disputes* dataset, allowing for this non-spatial dataset to be used in geographic analyses. Incident data on violence has recently become available also from military sources. The SIGACTS database (Berman, Shapiro and Felter 2008) has become an extremely valuable resource in analyses of recent violence in Iraq and Afghanistan (Lyall 2010*b*;*a*).

Geographic incident databases are also being created for other, less organized or less severe forms of violence or political contention. Researchers at the University of North Texas have recently published the Social Conflict in Africa (SCAD) database (Hendrix et al. 2010). Following an approach similar to the incident datasets described above, SCAD lists events of riots, strikes, protests, coups, and communal violence, each with spatial and temporal coordinates.

Demographic Demographic information can be spatially referenced in different ways. Raster datasets are sometimes used to capture the distribution of people across countries, as for example in the *Gridded Population of the World* (GPW) dataset (Center for International Earth Science Information Network CIESIN and International Center for Tropical Agriculture CIAT 2005). GPW is a raster with a resolution of 2.5 arc-minutes, which corresponds to about 5 km at the equator. For each raster cell, GPW gives the estimated number of people. Estimates exist for different years, starting in 1995 until 2010. Alternatively, a finer grid (the *GRUMP* dataset) is available from the same producer. The second major population raster is *LandScan* (Dobson et al. 2003). LandScan uses a finer resolution (30 arc-second, about 1 km), but unlike GPW, it is freely available only for certain categories of users and does not openly disclose its procedures for data generation. Based on these population datasets, it is possible to derive related rasters, for example on the extent of urban areas (also available from the

GPW website).

Various data sources on the spatial distribution of ethnic groups have recently been produced. The first of these attempts is the *Geo-referencing of Ethnic Groups* (GREG) project (Weidmann, Rød and Cederman 2010), a GIS dataset of ethnic group settlement regions across the globe. While other data collections on ethnic groups such as the Minorities at Risk dataset (Gurr 1993) include few spatial variables, GREG makes it possible to rely on GIS techniques to derive a variety of spatial and non-spatial indicators (Buhaug, Cederman and Rød 2008). Some examples described below demonstrate how this can be done. GREG captures the ethnic configuration of countries as a single static layer, with no changes over time. A subsequent project has attempted to produce a more dynamic spatial coding of group settlement regions. *GeoEPR* (Wucherpfennig et al. 2011) is a spatial extension to the *Ethnic Power Relations* dataset (Wimmer, Cederman and Min 2009). In GeoEPR, group polygons have a "lifespan", which makes it possible to represent changes in group settlements over time. However, despite the huge potential these datasets have for macro-level research on ethnicity, the coarse representation of group regions as separate polygons makes it difficult for example to capture varying degrees of ethnic mixing, as well as micro-level patterns of ethnic distribution such as ethnic neighborhoods in city.

Socio-economic A variety of GIS datasets map socio-economic variables. *G-ECON* is a raster dataset that provides estimates of sub-national economic activity at the grid cell level for 1 degree grid cells (Nordhaus 2006), based on country specific sources and estimates, with a temporal range from 1990 until 2000. A geographic dataset on poverty and infant mortality is being produced by the *Poverty Mapping Project* at Columbia University's Earth Institute (Storeygard et al. 2008). The data are available both in vector- and raster formats, and are referenced to subnational administrative units. An alternative way to measure economic activity is the use of satellite images of nighttime light emissions. A variety of products from different satellites are available, one of which is the *Nighttime Lights of the World* dataset released by National Oceanic and Atmospheric Administration's National Geophysical Data Center (`http://sabr.ngdc.noaa.gov/ntl/`). Agnew et al. (2008) use these data to identify the quality of everyday life in Baghdad and the impact of the US military surge. However, because of differential rates of electrification between developed and developing countries, these datasets have to be used with caution for cross-national studies. Doll (2008) provides a good overview of the advantages and problems with nighttime lights data.

Relatively good data exist on transportation networks and accessibility. ESRI, the producer of one of the most popular GIS software packages, provides a global dataset of major roads along with its *Digital Chart of the World* (ESRI 2006). Using this and other related datasets, researchers have produced a global map of accessibility. For small raster cells, this dataset gives the estimated travel time to the nearest major city, using common means of transportation (Uchida and Nelson 2010). Starr (2009), for example, use these data to measure the accessibility and "ease of interaction" across borders, and explores its possible relationship to conflict and cooperation.

Environmental Since many environmental variables can be measured using satellite imagery, there are a host of raster datasets available for this domain. The *GTOPO30* dataset is a global raster data on territorial elevation, developed by the United States Geological Survey, where elevation is measured at the level of grid cells with a resolution of 30 arc-seconds (`http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30_info`). Estimates of rainfall and related variables are provided in raster formats by the *Global Precipitation Climatology Project*, available at `http://cics.umd.edu/GPCP`. Existing research suggests that data on meteorological conditions can help gauge truly exogenous shocks influencing production in economies dominated by rain-feed agriculture and thus potentially improve causal identification of the effects of growth on conflict (Miguel, Satyanath and Sergenti 2004; Jensen and Gleditsch 2009; Ciccone 2011). Another very comprehensive collection of meteorological datasets is the *Global Precipitation Climatology Centre* administered by the German Meteorological service (`http://gpcc.dwd.de/`).

Types of land cover can also be approximated using satellite data, and are available as raster datasets. One of the most recent efforts to do this is the GlobCover project (Bontemps, Defourny and Van Bogaert 2010) that provides categories of land cover (cropland, forest, grassland, populated places) for 300 m cells globally (`http://ionia1.esrin.esa.int/`). GlobCover maps global land cover in regular intervals, so that the dataset can also be used to track changes occurring, for example, as a result of social or political processes.

Combined collections Since disaggregated spatial data come in different resolutions, combining different data sources will usually require scaling to some common resolution. As such, it may be helpful for many researchers to have a standardized grid structure that integrates different data sources to a common set of geographical cells. One example here is the *PRIO-GRID* project (Tollefsen, Strand and Buhaug ND), which provides a global grid with a resolution of 0.5 decimal degrees (roughly 50 km at the Equator). This dives the earth's landmass into approximately 62,000 cells in a cross-section (the PRIO-GRID data are annual and currently available for 1946-2008). Buhaug et al. (2011) for example, use PRIO-GRID to examine a number of hypotheses on how geographically varying local economic characteristics influence the likelihood that particular cells or locations will see the initial onset event in a conflict, controlling for other social and political factors believed to influence the risk of conflict.

Creating Spatial Datasets

Rather than relying on readily provided spatial datasets such as the ones described above, there are different procedures by which researchers can create new spatial datasets for their projects. As we have discussed above, in a spatial dataset the individual observations—either vector features or raster cells—have information about their geographic location attached to them. Therefore, creating a spatial dataset essentially means that we can record these spatial coordinates directly when the data is collected, or append spatial information to existing, non-spatial data, a step that is usually called "geo-referencing".

One way to directly collect spatial coordinates for observations is by means of the *Global Positioning System* (GPS). A GPS receiver is an electronic device that with the help of satellites can determine its geographic position on the globe with a high level of accuracy. Survey researchers typically take advantage of GPS by recording the geographic position at which the respondent lives, or where an interview was conducted. The *Demographic and Health Surveys* project, for example, which conducts surveys on various living standards and health related outcomes at the level of households helpful for evaluation population and health programs, routinely attaches GPS coordinates to their recorded data (`http://`

`www.measuredhs.com`). Using other GIS layers, it becomes possible to link the individual surveys to the particular context a respondent lives in, and amend the data collected in the survey using these layers. Hegre, Østby and Raleigh (2009) use these data to approximate geographical variation in poverty in Liberia. Data on violent incidents from military sources are usually tagged in the same way, with reporting military unit submitting GPS coordinates along with an incident report.

Existing, non-spatial datasets can be made GIS-compatible in different ways. Event datasets such as the ones described above that are created from news reports usually obtain spatial coordinates by converting place names into geographic coordinates. Media reports about incidents typically mention the village or city in which an incident happened. The location of these villages can then be found with the help of gazetteers. A gazetteer is a list of place names and their spatial coordinates, and most of them are now available online. Useful examples include the *Falling Rain* database (`http://www.fallingrain.com/world/index.html`), or the *GEOnet* Names Server at the NGA (`http://earth-info.nga.mil/gns/html/index.html`). Since English spellings of foreign place names rarely are standardized, the *JRC Fuzzy Gazetteer* (`http://dma.jrc.it/services/fuzzyg/`) is particularly convenient, since it retrieves place names even if the spelling does not match perfectly.

Alternatively, GIS databases can be created by converting existing maps into GIS-compatible formats. Vanzo (1999), for example, geocodes historical maps reflecting boundary changes to examine to what extent post-conflict borders reflect a tendency towards greater territorial compactness. Geocoding was also used to create the Weidmann, Rød and Cederman (2010) GREG dataset on ethnic groups described above. The first step in this process is the scanning of the map, if it does not yet exist in electronic form. Next, the map needs to be aligned correctly with the spatial reference system used by the GIS. Last, the spatial features of interest (e.g. settlement regions of ethnic groups) must be extracted from these maps. This can be done by tracing these features manually, or by applying a feature recognition algorithm to the scanned map. Various techniques exist to facilitate this process, and we refer the reader to the technical literature (Longley et al. 2010).

Approaches to GIS Analysis

The purpose of this section is to review different approaches to how GIS data and techniques have been used in international studies. We distinguish between spatial procedures to (i) create non-spatial data, (ii) generate static spatial variables, and (iii) study dynamic spatial processes.

Spatial Procedures to Create Non-Spatial Data

Even when the original research question is not inherently spatial or geographically focused, spatial data can often be helpful to create more specific and detailed information on the key quantities of interest, as the relevant features vary spatially. As an example, we demonstrate how spatial data can help improve information on the effects of intergroup inequality and civil war.

There is a long research tradition in examining the role of grievances in the sense of economic inequality and their possible effects on conflict (Gurr 1970; Lichbach 1989). Although earlier research tended to find support for more political protest under higher land inequality or income inequality (Muller and Seligson

1987), many more recent studies of civil war claim to present empirical conclusive evidence against this thesis (Collier and Hoeffler 2004), as data there is no clear relationship between conventional measures of interpersonal income inequality and conflict. However, Cederman, Weidmann and Gleditsch (2011) argue that theories of conflict do not suggest that greater social inequality should lead to violence, but rather that inequalities that coincide with other salient cleavages such as ethnic divisions are likely to exacerbate tension and spur violent mobilization (see also Stewart 2008; Østby 2008).

Spatial data can help us generate measures of these "horizontal inequalities". Since most ethnic groups tend to be spatially confined, the approach is to combine spatial data on ethnic group settlements with disaggregated economic data. Figure 3 illustrates the construction of the measures for Yugoslavia. We use the GECON data, which contain information on economic activity by geographic grid cells, and the GeoEPR data, which provide polygon data for group settlement areas. Overlaying these will give us a measure of total economic activity by group settlement area. We can then consider group inequality by comparing per-capita wealth for each group with the national average, with values above 1 for relatively more affluent groups and values below 1 for poorer groups (Figure 3, right). As expected, the ratios indicate that Albanians in Kosovo are on average poorer than the national average, while the Croats and the Slovenes are wealthier. Using such group inequality ratios in a large global sample, Cederman, Weidmann and Gleditsch (2011) find that large inequalities that follow ethnic lines are associated with a higher risk of ethnic conflict, both when groups are relatively disadvantaged and when groups are privileged.

Static Spatial Disaggregation

In the previous section, we discussed how spatially varying features can be used to create proxies or measures for other characteristics of interest. However, the spatial variation itself is often of direct interest to the research, especially when certain characteristics or outcomes of interest to the researcher vary within large units, for example, countries or states. As an example, we consider how the information on the specific locations of conflict and violence can help inform research on the causes and consequences of war and evaluate contending hypotheses.

Much of the existing research on civil war have considered "civil war" as a dichotomous outcome of the country level, where states are either "at war" or not over some specific period. However, civil wars rarely engulf entire countries and come in many different degrees, both in terms of the severity of the fighting as well as the geographical scope. The *Conflict Sites* dataset expands the Uppsala Armed Conflict Data to include a geographical representation of the conflict zone where violence takes place, using a polygon representation of conflict. Figure 4 displays two important examples of variation in the distribution in civil wars. For example, the Chechen War in 1995 (left panel) is confined to a relatively small and peripheral part of the territory of Russia, and clearly does not extend to the entire country. Since the conflict zone is small and unlikely to influence national figures, looking for consequences of the conflict in national level data may overlook the actual local impact of the conflict. Moreover, if the region of Chechya is atypical and divergent from the rest of the Russia evaluating hypotheses on the causes of conflict with national level data may be inappropriate and misleading. Conversely, the case of Democratic Republic of Congo in 2007 (right panel) provides an example of a country experiencing two distinct civil wars that take place in completely different parts of the country. In the East, the government fights the National Congress for the Defence of the People (CNDP), which a Tutsi dominated organization, which rejects the authority of the government and arguably seeks autonomy or secesion. In the West, the government fights the Bundu Dia Kongo, which claiming to represent the Kongo people and seek to restore pre-colonial boundaries. Treating the country at large as "at war" distracts our attention from the distinct actors and conflictual interactions taking place.

Spatial data provide opportunities for more systematic analyses of how spatially varying features influence the causes of conflict. If all politics indeed are local, then the causes of the conflict are more likely to observable or reflected in the characteristics of the areas where they occur rather than features of state at large. By construction, many conventional country-level statistics such as Gross Domestic Product per capita or ethnic fractionalization are averages that reflect population density, and will not reflect variation within countries. Buhaug and Lujala (2005) examine the characteristics of conflict zones and compare these to other areas within the same country without conflict. By comparing (GIS) data for other space-varying factors plausible relevant to conflict onset, at the scale of the conflict, they demonstrate that conflict zones tend to be very different from other areas of a country. As such, analyses that compare country characteristics based on average or aggregate measures to conflict may miss out on important spatial variation within countries. Since civil wars tend to be fought by small groups, often in thinly populated peripheral areas, the risk of conflict may be better reflected by "worst case" indicators, or measures of the geographical areas most likely to see conflict, rather than population weighted measures. More generally, researchers should think carefully about correspondence between actual measures and the underlying theoretical concept. Just because a particular measure is available or a conventional measure in existing research does not necessarily mean that it is a suitable indicator for testing a particular argument.

Buhaug and Gates (2002) examine what factors may account for variation in the geographical scope of conflict. They find strong evidence that the presence of natural resources within conflicts and their overall duration influence the geographical scope of conflicts, and their results suggest a possible endogenous relationship between geographical scope and the peripheral location of conflict.

The consequences of conflict are likely to be proportional to their magnitude. Although civil wars can be shown to have a negative impact on social and economic development, it seems unreasonable to expect that conflicts with a limited geographical scope would have identical consequences to large conflicts with broad geographical reach. Although there has been a great deal of interest in how conflict affects other outcome such as development and health (Collier et al. 2003; Ghobarah, Huth and Russett 2003), we are not aware of research to date examining how the geographical scope or intensity of conflict influences the outcomes or consequences of conflict.

Dynamic Spatial Processes

So far we have reviewed example of static spatial data, or cases where researchers are interested in assessing how some feature varies across space at some given point in time. For other research questions, the dynamics of change over time may be the core issue of interest, and spatial position may itself be seen as causally important. We consider the diffusion of conflict as an example of a dynamic spatial process. There are arguably many important mechanisms that can create spatial dependence between different actors, units, or locations and increase the risk of conflict. More generally, if ongoing conflict in one country can affect the risk of conflict in other states, the individual conflict outbreak are not independent as the outcomes are shaped by events and outcomes in other, connected observations.

Much of the conventional research on civil war has adopted a "closed polity" approach to the study of civil war, assuming that the relevant causes of internal conflicts must be found within the boundaries of the country experiencing conflict (Gleditsch 2007). However, there are strong theoretical reasons why the risk of conflict may be shaped by events and features in other states, especially neighboring countries. For example, many conflicts often involve demands for autonomy or independence from particular ethnic communities, who often reside in multiple countries (Cederman, Girardin and Gleditsch 2009; Horowitz 1985; Lake and Rothchild 1998). The decision to contest the state militarily can in this case be influenced by experiences of the group in another state, or the ability to rely on financial or military support from kin in another state. An ongoing civil war in a neighboring country can increase the availability of arms and recruits and make it relatively less costly to mobilize insurgencies (Lischer 2005; Salehyan and Gleditsch 2006). Hostile relations between states can give governments incentives to support insurgencies in a neighboring state to undermine their rival (Davis and Moore 1997; Salehyan, Gleditsch and Cunningham 2011).

If civil war in one state indeed increases the risk of seeing conflict in another state, then we should be able to see evidence of a spatial process using GIS data, where conflict affected areas cluster geographically and evolve in particular ways over time. The Great Lakes Region of Africa in the 1990s is often cited as an example of a cluster of interdependent civil wars (Prunier 2008; McNulty 1999). As such, it is instructive to consider how the spatio-temporal evolution of conflicts in this region is reflected in the *Conflict Sites* dataset. Figure 5 displays the conflict polygons for the years 1991 to 1999, for a section of the eastern part of the Democratic Republic of Congo. In 1991, the civil war in Rwanda first erupted when the Rwandan Patriotic Front (RPF) invaded from Uganda, where a Tutsi refugee population of about 200,000 individuals had organized militarily, with assistance from Ugandan authorities and partial integration of rebel forces in the regular army. The conflict polygon in Rwanda is clearly located on the border, reflecting the important role of the ties to Uganda in the conflict, where the RPF on occasion retreated into Uganda to regroup and rearm during the initial years of the conflict.

The maps for the subsequent years reflect the ongoing civil war in Rwanda, and the eventual escalation to encompass the whole country around the Rwandan genocide in 1994 and the eventual RPF victory, where the RPF leader Kagame became the new President of Rwanda. The Rwandan civil war generated to a major refugee crisis, where Hutu refugees fled to neighboring countries, in particular Zaire. These refugee camps in turn provided a fertile environment for a Hutu insurgent movement, the Rassemblement Démocratique pour le Rwanda. In response, the Kagame government in Rwanda turned to provide support for an insurgent group in Zaire, the Alliance of Democratic Forces for the Liberation of Congo (AFDL) led by Kabila. This is reflected in the conflict polygon in Eastern Zaire in 1996, again clearly on the border with Rwanda. The maps for 1997 show the subsequent escalation of the conflict in Zaire, where the AFDL overthrows Mobutu in 1997 and declares a new Democratic Republic of the Congo (DRC).

However, the maps for the subsequent years demonstrate how peace has remained elusive in the region, and possibly may have fuelled an escalation of the civil war in Uganda.

Inspecting such spatial representations of the data can be helpful for assessing the plausibility of claims about the interdependence of different conflicts and the importance of borders, but obviously does constitute a particularly rigorous test, especially since there may be many countries bordering states with civil war that do not experience conflict. A number of studies have used techniques from spatial statistics to test the importance of spatial proximity to conflict more systematically and found considerable support that this increases the risk of conflict (Bosker and de Ree 2011; Gleditsch 2007; Ward and Gleditsch 2002). In their sensitivity analysis of civil war, Hegre and Sambanis (2006) report that this is one of the key features that appear to have a robust positive influence on the risk of civil war. Other researchers have estimated the effects of specific mechanisms or transnational linkages, including transborder ethnic kin (Bosker and de Ree 2011; Cederman, Girardin and Gleditsch 2009), or refugee flows (Salehyan and Gleditsch 2006).

Despite the extensive academic interest in the diffusion of conflict, most of these studies have relied on data at the country level rather than GIS data. Geocoded data provides the opportunity to develop more fine grained measures of the spatial distribution of the relevant features, such as conflict zones. For example, while the conflict in Chechnya arguable can be expected to increase the risk of instability among Russia's neighbors in the Caucasus, it seems unreasonable to expect that neighboring states far from the conflict zone such as Finland or Norway should see an elevated risk of conflict. Buhaug and Gleditsch (2008) examine to what extent the risk of conflict contagion depend on exposure to neighboring conflict versus ethnic ties to the state experiencing conflict, and find evidence suggesting that the latter tends to dominate. However, these analyses were based on data that are now partly outdated, and there are many opportunities for more refined analyses using currently spatial data and resources.

More recent studies have looked at diffusion within individual conflicts. Schutte and Weidmann (2011) distinguish between two types of diffusion, relocation and escalation, each of which is the result of a particular type of warfare. They find that civil wars primarily exhibit ecalation diffusion as a result of irregular warfare without conventional front lines. Weidmann and Ward (2010) consider a model of spatial and temporal diffusion for conflict events in Bosnia for monthly event counts from March 1992 to October 1995. Their results demonstrate that there is a strong spatial as well as temporal dimension to the violence in the Bosnia conflict, where violence is likely to recur over time and spread spatially. Moreover, they show that taking advantage of this information can substantial improve the ability to forecasts conflict over purely structural models.

Although our examples here have focused primarily on conflict, the general points that we have emphasized apply much more widely. Diffusion processes are not necessarily inherently spatial (Beck, Gleditsch and Beardsley 2006; Coleman 1964; Rogers 1962; Wejnert 2002), but it will often be the case that spatial position exerts a powerful influence shaping the network of interdependence between actors (Simmons and Elkins 2004; Simmons, Dobbin and Garrett 2008; Siverson and Starr 1991; Ward and Gleditsch 2008). Moreover, other forms of connections and network position can often be given spatial representations, and analogies to spatial data are often helpful for analyzing such data (Bavaud 1998; Guyon

1995).

Challenges in Spatial Analysis

Despite all its promises, spatial analysis and GIS data entail a series of challenges, the most important of which we discuss here.

Practical Challenges

Given the obvious potential for applications of GIS and spatial data for international studies, it may seem surprising to many that there has not been more use of these methods and techniques. On the one hand there are many important practical obstacles that have prevented more widespread use of GIS and spatial data. GIS and spatial data have been dominated by expensive proprietary software solutions, notably the ESRI ArcGIS package. Since ESRI has had a strong market leader position and much of the available training has focused on these packages, the prohibitive costs have in all likelihood created significant barriers to more widespread use. Beyond the costs of proprietary software, there are many technical challenges in working with spatial data. The lack of training in spatial methods that currently exists in many social science curricula imposes high entry costs for researchers.

These practical obstacles are likely to be overcome by the recent emergence of open-source software tools for spatial data. In particular, there has been considerable development of spatial extensions in the open source package *R*, documented with examples in Bivand, Pebesma and Gómez-Rubio (2008). Since many social scientists are already familiar with R, the ability to work with spatial data generation and statistical analysis within the same framework has helped disseminate more widespread use of GIS and spatial data.

Theoretical and Research Design Challenges

Using spatial data in inferential analyses involves important theoretical issues that we have so far glossed over. In particular, researchers must consider the choice of units and appropriate resolution for analysis. Although disaggregation can certainly be helpful, more disaggregation is not inherently better and the appropriate units or analysis and level of aggregation is above all a theoretical question (Cederman and Gleditsch 2009).

In some cases, researchers examine very high levels of aggregation such as international systems (Kaplan 1957). Measuring system concepts such as "polarity" or "power-concentration" require a number of decisions as to how attributes of lower level units should be combined at the system level. Dyads, or pairs of states, are common units of analysis in studies of international conflict, and require decisions on how one should construct dyadic measures from the attributes of two states (i.e., minimum vs. average) or whether one should focus on directed or undirected behavior (Goertz 2005). In cross-national research, a useful contrast can be drawn between studies with the country as the unit of analysis and studies with individuals as units. The former choice may be appropriate for research questions where countries are the fundamental actors of interest, for example studies that examine the central government, the total size of the economy, or the "average" income or other typical characteristics in a country. By contrast, individual respondents may be helpful for research questions pertaining to decisions by individual actors, for example, why individuals join rebel movements and reintegration, or for studying variation across individuals (Kalyvas

2006; Humphreys and Weinstein 2008).

In between these two extremes, however, there are many other alternatives that may be useful for analysts. For some questions, illustrated above, focusing on ethnic groups and individual organizations may be more helpful. Civil war, for example, requires mobilization and collective action, and a focus on groups or organizations as the key units of analysis may be helpful for understanding agency in violent conflict. Spatial data sources can help approximate characteristics of their constituencies. In other cases, specific geographic areas such as sub-national region may be more appropriate to capture geographical variation and the locus of events and interactions. In many cases, it may be possible to combine information from different levels, and scale information collected at different levels "up" or "down" to other units of analysis. However, the appropriate resolution or choice of units of analysis is always ultimately a theoretical question that must be defended in each case with reference to the research question.

Statistical Challenges in Spatial Analysis

Spatial dependence, or the notion that units of analysis may not be independent of one another or a random sample across space in a cross-section, has already been recognized as an important problem in cross-national research (Beck, Gleditsch and Beardsley 2006; Franzese and Hays 2008). There exists a large literature on spatial statistics (Cressie 1991; Ward and Gleditsch 2008), and a full overview of this literature is beyond the scope of this review. However, it must be emphasized that spatial disaggregation to smaller units in general will tend to increase the problem of spatial dependence. In essence, if features are spatially correlated, then increasingly more detailed nearby observations tend to be similar to one another. As such, additional observations contribute less independent observations and the apparent number of cases can be much larger than the "effective" number of cases.

Spatial statistical models for categorical or event count variables and panel data structure are much more complicated than for continuous variables and cross-sectional data and often rely on simulation based methods. Many efforts have looked at how spatial models from other fields such as ecology and image restoration can help contribute to social science applications (Ward and Gleditsch 2002; Weidmann and Ward 2010), and there is likely to be considerable future research in this area.

There is an important conceptual distinction between research questions that try to examine the effect of some particular covariate and applications that try to model spatial diffusion processes. Although the latter can be a very complex task, if the interest lies exclusively in the former, then it may not be necessary to examine a full population, and and ideas from the subsampling literature and case-control designs where outcomes of the event of interest are compared with a suitable control sample may be helpful (King and Zeng 2001; Heagerty, Ward and Gleditsch 2002). For example, since logit is invariant to changes in the marginal totals, comparing cases with the events of interest with a smaller representative sample of "non-events" will not affect estimates of the coefficient for specific covariates, although the intercept will need to be adjusted. Buhaug et al. (2011), for example, examine so-called variograms for cell data to examine how the spatial covariance betweens observation depends on the distance between them in order to assess possible threshold beyond which observations can be considered independent of one another.

Conclusion

For social scientists, GIS has tended to be equated with generating maps. Although maps are important and much can be learned from visualizing data on a map, we see the ability to generate new measures from spatial data and modeling spatial processes as the key promise of GIS in international studies.

We started this review by a discussion of the pioneering research by Richardson on geography and conflict, which was mainly done by hand, and argued that modern GIS and the increasing availability of spatial data provide many opportunities for advancing this line of research. There is an interesting analogy here to Richardson's (1922) work on weather forecasting, which proposed a system based on solving differential equations that was not computationally feasible at the time. Subsequent advances in computing, however, vindicated Richardson's ideas (Lynch 2006). The first modern computer ENIAC generated a weather forecast in 1950, and similar models are today used extensively for weather forecasting and modeling climate change.

We believe that our review provides strong support for our claim that use of GIS and spatial data has helped advance research along the lines of Richardson's original question. Spatial data have helped facilitate new approaches to the study of inequality and conflict, which at least to some provide a vindication of the role of grievances in civil war often dismissed in other research. The "closed polity" assumption seems untenable in studies of civil war, and we have learned important things about the risks of conflict diffusion, as well as specific conditions where conflicts are more or less likely to generate instability in other countries. Although this is an emerging area and much still remains to be done, we believe that Richardson would have been very pleased to see the results of existing research using GIS and spatial data.

The spatial perspective is not just a question of tools and technique, but also helps foster a substantively novel theoretical approach for understanding political events and outcomes. Whereas much research often takes states as predetermined units with fixed boundaries, Richardson's work on borders alerts us to the endogenous nature of borders, and how present day borders reflect historical and political processes that have generated and preserved borders. Whereas much comparative research traditionally may have treated individual countries as independent units, there is a great deal of new interest in the role of interdependence in a globalizing world.

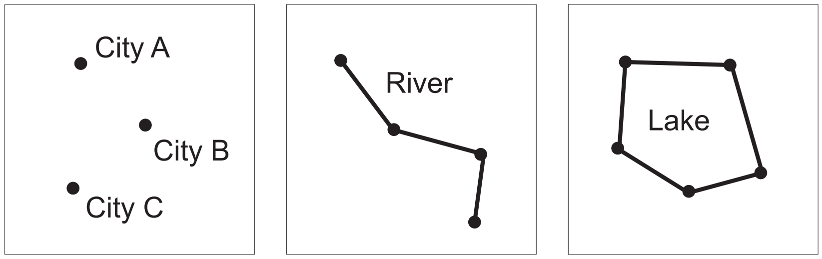

Figure 1: GIS Vector Data Types: Points, Lines and Polygons

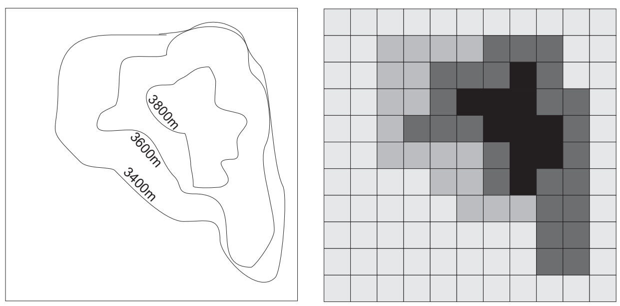

Figure 2: GIS Raster Data on Territorial Elevation

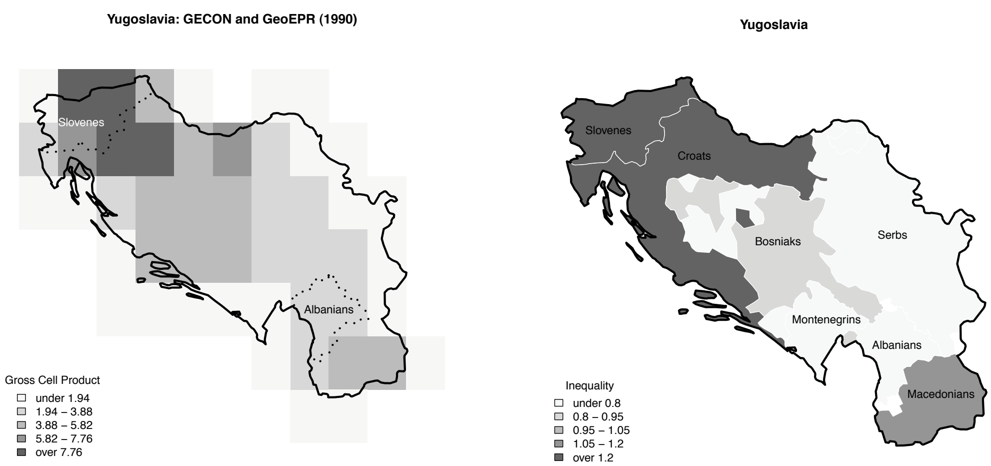

Figure 3: Example for the computation of the group wealth indicator. The G-ECON dataset on economic performance is overlaid with the group settlement regions from Geo-EPR (left). Aggregated the (partial) G-ECON cell values by group results in wealth estimates at the group level (shown as proportions of the national average, right).

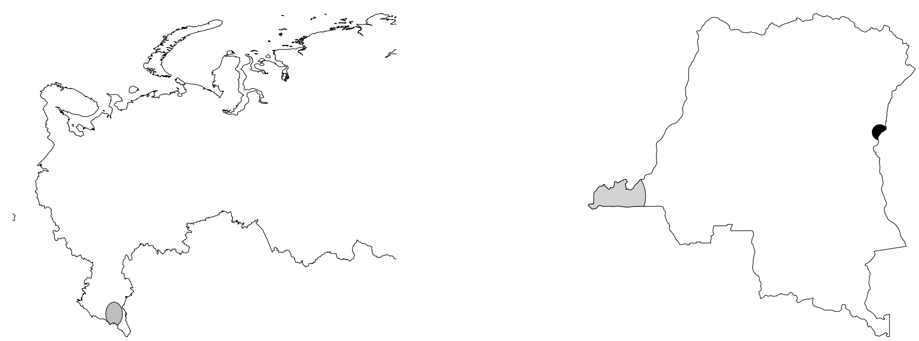

Figure 4: Example of conflict polygons from the `Conflict Sites` data, for the Chechen war in Russia in 1995 (left), and the 2007 civil wars in the Democratic Republic of Congo, where the government fights the National Congress for the Defence of the People (CNDP) in the East and the Bundu Dia Kongo (BDK) in the West (right).

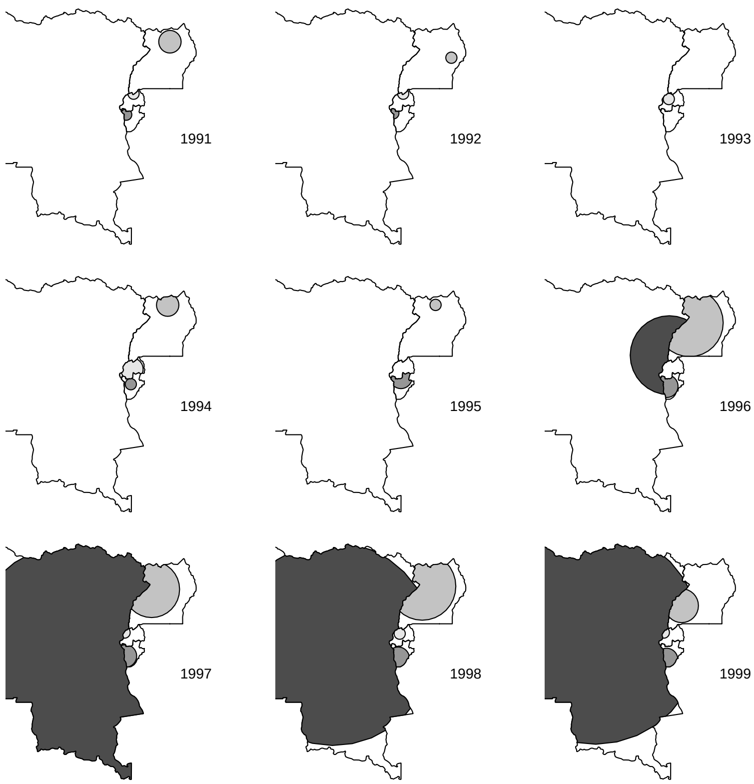

Figure 5: Conflict polygons in the Great Lakes region of Africa, 1991-99

References

Agnew, John, Thomas W. Gillespie, Jorge Gonzalez and Brian Min. 2008. "Baghdad Nights: Evaluating the US Military Surge Using Night Light Signatures." *Environment and Planning A* 40(10):2285–2295.

Alesina, Alberto, William Easterly and Janina Matuszeski. 2011. "Artificial States." *Journal of the European Economic Association* 9(2):246–277.

Bavaud, Francois. 1998. "Models for Spatial Weights: A Systematic Look." *Geographical Analysis* 30:153–171.

Beck, Nathaniel L., Kristian Skrede Gleditsch and Kyle C. Beardsley. 2006. "Space is More than Geography: Using Spatial Econometrics in the Study of Political Economy." *International Studies Quarterly* 50(1):27–44.

Berman, Eli, Jacob N. Shapiro and Joseph H. Felter. 2008. "Can Hearts and Minds Be Bought? The Economics of Counterinsurgency in Iraq.". NBER Working Paper No. 14606.

Bivand, Roger S., Edzer J. Pebesma and Virgilio Gómez-Rubio. 2008. *Applied Spatial Data Analysis with R.* New York: Springer.

Bontemps, Sophie, Pierre Defourny and Eric Van Bogaert. 2010. "GlobCover 2009 Product Description Manual." Technical report, European Space Agency and Catholic University of Leuven.

Bosker, Maarten and Joppe de Ree. 2011. "Ethnicity and the spread of civil war." Center for Economic Poliy Research Discussion Paper 8055.

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

Brody, Howard, Michael Russell Rip, Peter Vinten-Johansen, Nigel Paneth and

Stephen Rachman. 2000. "Map-making and myth-making in Broad Street: The London cholera epidemic, 1854." *The Lancet* 356:64–68.

Buhaug, Halvard and Jan Ketil Rød. 2006. "Local Determinants of African Civil Wars, 1970-2001." *Political Geography* 25(3):315–335.

Buhaug, Halvard and Kristian Skrede Gleditsch. 2008. "Contagion or Confusion? Why Conflicts Cluster in Space." *International Studies Quarterly* 52(2):215–233.

Buhaug, Halvard, Kristian Skrede Gleditsch, Helge Holtermann, Gudrun Østby and Andreas Forø Tollefsen. 2011. "It's the Local Economy, Stupid! Geographic Wealth Dispersion and Conflict Outbreak Location." *Journal of Conflict Resolution* tba(tba):forthcoming.

Buhaug, Halvard, Lars-Erik Cederman and Jan Ketil Rød. 2008. "Disaggregating Ethnic Conflict: A Dyadic Model of Exclusion Theory." *International Organization* 62(3):531–551.

Buhaug, Halvard and Päivi Lujala. 2005. "Accounting for Scale: Measuring Geography in Quantitative Studies of Civil War." *Political Geography* 24(4):399–418.

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

Cederman, Lars Erik and Kristian Skrede Gleditsch. 2009. "Special Issue on 'Disaggregating Civil War'." *Journal of Conflict Resolution* 53(4):487–495.

Cederman, Lars-Erik, Luc Girardin and Kristian Skrede Gleditsch. 2009. "Ethnonationalist Triads: Assessing the Influence of Kin Groups on Civil Wars." *World Politics* 61(3):403–437.

Cederman, Lars-Erik, Nils Weidmann and Kristian Skrede Gleditsch. 2011. "Horizontal Inequalities and Ethno-nationalist Civil War: A Global Comparison." *American Political Science Review* 105(2):forthcoming.

Center for International Earth Science Information Network CIESIN and International Center for Tropical Agriculture CIAT. 2005. "Gridded Population of the World v3 (GPWv3).". Available at http://sedac.ciesin.columbia.edu/gpw/.

Ciccone, Antonio. 2011. "Economic Shocks and Civil Conflict: A Comment." *American Economic Journal: Applied Economics* tba(tba):forthcoming.

Coleman, James S. 1964. *Introduction to Mathematical Sociology.* Glencoe, IL: Free Press.

Collier, Paul and Anke Hoeffler. 2004. "Greed and Grievance in Civil War." *Oxford Economic Papers* 56:663–595.

Collier, Paul, Lani Elliott, Håvard Hegre, Anke Hoeffler, Marta Reynal-Querol and Nicholas Sambanis. 2003. *Breaking the Conflict Trap: Civil War and Development Policy.* Washington, DC: World Bank & Oxford University Press.

Cox, Kevin R., Murray Low and Jennifer Robinson, eds. 2007. *The Sage Handbook of Political Geography.* Thousand Oaks, CA: Sage.

Cressie, Noel A. C. 1991. *Statistics for Spatial Data.* Chichester: Wiley.

Cunningham, Kathleen Gallagher and Nils B. Weidmann. 2010. "Shared Space: Ethnic Groups, State Accommodation and Localized Conflict." *International Studies Quarterly* 54(4):1035–1054.

Davis, David R. and Will H. Moore. 1997. "Ethnicity Matters: Transnational Ethnic Alliances and Foreign Policy Behavior." *International Studies Quarterly* 41(1):171–184.

Deiwiks, Christa. 2010. "The Curse of Ethno-federalism? Ethnic Group Regions, Subnational Boundaries and Secessionist Conflict." Paper prepared for the Annual Convention of the International Studies Association.

Dobson, J. E., E. A. Bright, P. R. Coleman and B. L. Bhaduri. 2003. LandScan: A Global Population Database for Estimating Population at Risk. In *Remotely Sensed Cities*, ed. V. Mesev. London: Taylor and Francis.

Doll, Christopher N.H. 2008. "CIESIN Thematic Guide to Night-time Light Remote Sensing and its Applications." Center for International Earth Science Information Network of Columbia University. Available at http://sedac.ciesin.columbia.edu/tg/.

ESRI. 2006. "Data and Maps collection for ArcGIS.". Available at http://www.esri.com/data/data-maps/overview.html.

Ethington, Philip J. and Jason A. McDaniel. 2007. "Political Places and Institutional Spaces: The Intersection of Political Science and Political Geography." *Annual Reviews of Political Science* 10:207–242.

Fearon, James D. and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 91(1):75–90.

Flint, Colin and Peter J. Taylor. 2007. *Political Geography: World-Economy, Nation-State and Locality, 5th Edition*. Harlow, Essex: Longman, Scientific and Technical.

Franzese, Robert J. and Jude C. Hays. 2008. "Interdependence in Comparative Politics: Substance, Theory, Empirics, Substance." 41(4-5):742–780.

Ghobarah, Hazem Adam, Paul Huth and Bruce M. Russett. 2003. "Civil Wars

Kill and Maim People-Long after the Shooting Stops." *American Political Science Review* 97(2):189–202.

Gleditsch, Kristian Skrede. 2007. "Transnational Dimensions of Civil War." *Journal of Peace Research* 44(3):293–309.

Gleditsch, Kristian Skrede and Michael D. Ward. 1999. "Interstate System Membership: A Revised List of Independent States since 1816." *International Interactions* 25:393–341.

Gleditsch, Kristian Skrede and Michael D. Ward. 2005. "Visualization in the Study of International Relations.".

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.

Goertz, Gary. 2005. *Social Science Concepts: A User's Guide.* Princeton, NJ: Princeton University Press.

Gurr, Ted R. 1970. *Why Men Rebel.* Princeton, NJ: Princeton University Press.

Gurr, Ted Robert. 1993. *Minorities at Risk: A Global View of Ethnopolitical Conflicts.* Washington, DC: United States Institute of Peace Press.

Guyon, Xavier. 1995. *Random Fields on a Network: Modeling, Statistics, and Applications.* New York: Springer.

Hallberg, Johan Dittrich. 2011. "Conflict Site 1989-2008: A Geo-Referenced Dataset on Armed Conflict." Typescript, Centre for the Study of Civil War.

Heagerty, Patrick, Michael D. Ward and Kristian Skrede Gleditsch. 2002. "Windows of Opportunity: Window Subseries Empirical Variance Estimators in International Relations." Political Analysis." 10:304–317.

Hegre, Håvard, Gudrun Østby and Clionadh Raleigh. 2009. "Poverty and Civil War Events: A Disaggregated Study of Liberia." *Journal of Conflict Resolution* 53(4):598–623.

Hegre, Håvard and Nicholas Sambanis. 2006. "A Sensitivity Analysis of the Empirical Literature on Civil War Onset." *Journal of Conflict Resolution* 50(4):508–535.

Hendrix, Cullen S., Idean Salehyan, Christina Case, Christopher Linebarger, Emily Stull and Jennifer Williams. 2010. "The Social Conflict in Africa Database: New Data and Applications." Working paper, the University of North Texas. Available at http://ccaps.strausscenter.org/scad/conflicts.

Hess, Gregory D. 1995. "An Introduction to Lewis Fry Richardson and His Mathematical Theory of War and Peace." *Conflict Management and Peace Science* 14(1):77–113.

Horowitz, Donald. 1985. *Ethnic Groups in Conflict*. Berkeley, CA: University of California Press.

Humphreys, Macartan and Jeremy M. Weinstein. 2008. "Who Fights? The Determinants of Participation in Civil War." *American Journal of Political Science* 52(2):436–455.

Jaggers, Keith and Ted R. Gurr. 1995. "Tracking Democracy's 'Third Wave' with the Polity III data." *Journal of Peace Research* 32(4):469–82.

Jensen, Peter Sandholt and Kristian Skrede Gleditsch. 2009. "Rain, Growth, and Civil War: The Importance of Location. Defence and Peace Economics." 20(5):359–372.

Kalyvas, Stathis. 2006. *The Logic of Violence in Civil War*. Cambridge: Cambridge University Press.

Kaplan, Morton A. 1957. *System and Process in International Politics*. New York: Wiley.

King, Gary and Langche Zeng. 2001. "Logistic Regression in Rare Events Data." *Political Analysis* 9(2):137–163.

Lake, David A. and Donald Rothchild, eds. 1998. *The International Spread of Ethnic Conflict: Fear, Diffusion, and Escalation*. Princeton, NJ: Princeton University Press.

Law, Gwillim. 1999. *Administrative Subdivisions of Countries*. Jefferson, NC: McFarland and Company.

Lichbach, Mark Irving. 1989. "An Evaluation of 'Does Economic Inequality Breed Political Conflict?' Studies." *World Politics* 41(4):431–470.

Lischer, Sarah Kenyon. 2005. *Dangerous Sanctuaries: Refugee Camps, Civil War, and the Dilemmas of Humanitarian Aid*. Cornell: Cornell University Press.

Longley, Paul A., Michael F. Goodchild, David J. Maguire and David W. Rhind. 2010. *Geographic Information Systems and Science*. 3rd ed. Chichester: Wiley.

Lyall, Jason. 2010*a*. "Death from Above: U.S. Airstrikes, Aid, and the Conditionality of Violence in Iraq, 2004-08." Typescript, Yale University.

Lyall, Jason. 2010*b*. "State Coercion and Civilian Victimization in Civil War: Evidence from Airstrikes in Afghanistan." Typescript, Yale University.

Lynch, Peter. 2006. *The Emergence of Numerical Weather Prediction: Richardson's Dream*. Cambridge: Cambridge University Press.

Mandelbrot, Benoit. 1967. "How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension." *Science* 156(3775):636–638.

McNulty, Mel. 1999. "The Collapse of Zaire: Implosion, Revolution, or External Sabotage?" *Journal of Modern African Studies* 39(1):53–82.

Melander, Erik and Ralph Sundberg. 2011. "Climate Change, Environmental Stress, and Violent Conflict: Tests introducing the UCDP Georeferenced Event Dataset." Typescript, Uppsala Conflict Data Program.

Miguel, Edward, Shanker Satyanath and Ernest Sergenti. 2004. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." *Journal of Political Economy* 112(4):725–753.

Morrow, James. 2000. The Ongoing Game-Theoretic Revolution. In *Handbook of War Studies II*, ed. Manus I. Midlarsky. Ann Arbor, MI: University of Michigan Press pp. 164–192.

Muller, Edward N. and Mitchell A. Seligson. 1987. "Inequality and Insurgency." *American Political Science Review* 87(2):425–451.

Nicholson, Michael. 1999. "Lewis Fry Richardson and the Study of the Causes of War." *British Journal of Political Science* 29(3):541–563.

Nordhaus, William D. 2006. "Geography and Macroeconomics: New Data and New Findings." *Proceedings of the National Academy of Sciences USA* 103(10):3510–3517.

Østby, Gudrun. 2008. "Polarization, Horizontal Inequalities and Violent Civil Conflict." *Journal of Peace Research* 45(2):143–162.

Patterson, Tom and Nathaniel V. Kelso. 2011. "The Natural Earth Project." Available at http://www.naturalearthdata.com.

Prunier, Gerard. 2008. *Africa's World War: Congo, the Rwandan Genocide, and the Making of a Continental Catastrophe.* Oxford: Oxford University Press.

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, David E. Cunningham, Lars Wilhelmsen and Nils Petter Gleditsch. 2006. "Conflict Sites 1946–2005." Codebook, available at http://www.prio.no/CSCW/Datasets/Armed-Conflict/Conflict-Site/.

Richardson, Lewis F. 1960*a*. *Arms and Insecurity.* Chicago and Pittsburgh, PA: Quadrangle/Boxwood.

Richardson, Lewis F. 1960*b*. *Statistics of Deadly Quarrels.* Chicago and Pittsburgh, PA: Quadrangle/Boxwood.

Richardson, Lewis F. 1992. The Problem of Contiguity: An Appendix to Statistics of Deadly Quarrels. In *Collected Papers of Lewis Fry Richardson*, ed. Oliver M. Ashford, P.G. Drazin, J.C.R. Hunt, Paul Smoker and Ian Sutherland. Cambridge, New York: Cambridge University Press.

Richardson, Lewis Frye. 1922. *Weather Prediction by Numerical Process.* Cambridge: Cambridge University Press.

Rogers, Everett M. 1962. *Diffusion of Innovations.* Glencoe, IL: Free Press.

Salehyan, Idean and Kristian Skrede Gleditsch. 2006. "Refugee Flows and the Spread of Civil War." *International Organization* 60(2):335–366.

Salehyan, Idean, Kristian Skrede Gleditsch and David Cunningham. 2011. "Explaining External Support for Insurgent Groups." *International Organization* TBA(TBA):TBA.

Schutte, Sebastian and Nils B. Weidmann. 2011. "Diffusion Patterns of Violence in Civil Wars." *Political Geography* XX(XX). Forthcoming.

Simmons, Beth A., Frank Dobbin and Geoffrey Garrett, eds. 2008. *The Global Diffusion of Markets and Democracy*. Cambridge: Cambridge University Press.

Simmons, Beth and Zachary Elkins. 2004. "The Globalization of Liberalization: Policy Diffusion in the International Political Economy." *American Political Science Review* 98(1):171–189.

Siverson, Randolph M. and Harvey Starr. 1991. *The Diffusion of War: A Study in Opportunity and Willingness*. Ann Arbor, MI: University of Michigan Press.

Snyder, Richard. 2001. "Scaling Down: The Subnational Comparative Method." *Studies in Comparative International Development* 36(1):93–110.

Starr, Harvey. 2009. "Borders, Ease of Interactions, Transactions and Cooperation: Tracking Integration in the EU Across Waves of Expansion." Typescript, University of South Carolina.

Stewart, Frances, ed. 2008. *In Horizontal Inequalities and Conflict: Understanding Group Violence in Multiethnic Societies*. Houndmills: Palgrave Macmillan.

Storeygard, Adam, Deborah Balk, Mark Levy and Glenn Deane. 2008. "The Global Distribution of Infant Mortality: A Subnational Spatial View." *Population, Space and Place* 14(3):209–229.

Tollefsen, Andreas Forø, Håvard Strand and Halvard Buhaug. ND. "PRIO-GRID: A Unified Spatial Data Structure." *Journal of Peace Research* tba(tba):forthcoming.

Uchida, Hirotsugu and Andrew Nelson. 2010. Agglomeration Index: Towards a

New Measure of Urban Concentration. In *Development in an Urban World*, ed. B. Guha-Khasnobis. UN-WIDER.

Vanzo, John P. 1999. Border Configuration and Conflict: Geographical Compactness as a Territorial Ambition of States. In *A Road Map to War: Territorial Dimensions of International Conflict*, ed. Paul F. Diehl. Nashville: Vanderbilt University Press.

von Clausewitz, Carl. 1832/1984. *On War.* Princeton, NJ: Princeton University Press.

Ward, Michael D. 2010. Statistical Analysis of International Interdependencies. In *The International Studies Encyclopaedic Compendium: Scientific Studies of International Processes, Volume 10.* Oxford: Wiley-Blackwell pp. 6615–6628.

Ward, Michael D. and Kristian Skrede Gleditsch. 2002. "Location, Location, Location: An MCMC Approach to Modeling the Spatial Context of War and Peace." *Political Analysis* 10(2):244–260.

Ward, Michael D. and Kristian Skrede Gleditsch. 2008. *Spatial Regression Models.* Thousand Oaks, CA: Sage.

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

Weidmann, Nils B., Jan Ketil Rød and Lars-Erik Cederman. 2010. "Representing Groups in Space: A New Dataset." *Journal of Peace Research* 47(4):491–499.

Weidmann, Nils B. and Michael D. Ward. 2010. "Predicting Conflict in Space and Time." *Journal of Conflict Resolution* 54(6):883–901.

Weidmann, Nils and Kristian Skrede Gleditsch. 2010. "Mapping and Measuring Country Shapes: The cshapes Package." *R Journal* 2(1). online at http://journal.r-project.org/archive/2010-1/.

Wejnert, Barbara. 2002. "Integrating Models of Diffusion of Innovations: A Conceptual Framework." *Annual Review of Sociology* 28:297–306.

Wimmer, Andreas, Lars-Erik Cederman and Brian Min. 2009. "Ethnic Politics and Armed Conflict: A Configurational Analysis of a New Global Dataset." *American Sociological Review* 74(2):316–337.

Wucherpfennig, Julian, Nils B. Weidmann, Lars-Erik Cederman and Andreas Wimmer. 2011. "Politically Relevant Ethnic Groups across Space and Time: Introducing the GeoEPR Dataset." *Conflict Management and Peace Science* XX(XX). Forthcoming.