Spatial Analysis in Political Geography

Andrew Linke and John O'Loughlin

Forthcoming in **A Companion to Political Geography** 2^{nd} edition Wiley-Blackwell, 2016
Edited by J. Agnew, V. Mamadouh, A. Secor and J. Sharp

Abstract (currently 153 words): *Two fundamental concepts in spatial analysis resonate both in political geography and other disciplines that examine geographic distributions of human phenomena. The meanings and significance of context (place) and nonstationarity (spatial autocorrelation) have to date been examined haphazardly. Fundamental questions relating to the definition of neighborhood (and the related "Modifiable Areal Unit Problem") remain only partly answered. Similarly, key and inconsistent choices in the handling of spatial autocorrelation (including choices of spatial weights for analysis) are as important as four decades ago when systematic examination of these issues began. Three recent developments – increased data availability, data integration across scales, and new analytical tools – have shifted the focus of spatial analysis in political geography. After reviewing these welcome improvements, an illustration through a stylized example of merging aggregate violent events and public opinion survey data from a representative population sample of the North Caucasus of Russia demonstrates spatial analytical opportunities for contemporary political geography.*

Authors: Andrew M. Linke, John O'Loughlin

Andrew M. Linke bio note (currently 64 words): Postdoctoral Research Associate at University of Colorado-Boulder Institute of Behavioral Science. PhD 2013 from the University of Colorado-Boulder Department of Geography. Research interests in political violence, political geography, climate change and conflict, African politics, spatial statistics and geographic information systems. Research articles in *Political Geography, International Interactions, Journal of Peace Research, Proceedings of the National Academy of Sciences, Eurasian Geography and Economics*, among others.

John O'Loughlin bio note (currently 71 words): Professor of Geography at University of Colorado-Boulder**.** PhD 1973 from the Pennsylvania State University. Editor of *Political Geography* since 1981 and *Eurasian Geography and Economics* since 2001. Research interests in the geographies of conflict including the relationship between climate change and conflict in Sub-Saharan Africa and post-Soviet conflicts. Also published on electoral geography of Nazi Germany, regional political geographies of Ukraine, geopolitical views of Russians, and social capital in post-Communist societies.

Keywords: Geographic Information Systems, Context, Place, Quantitative Methods, Russia, North Caucasus, Violent Conflict, Survey Data, Integrating Individual and Aggregate Data

Body (follows, currently 7652 words including references, headings, and captions):

I. INTRODUCTION

In contemporary spatial and quantitative study of political geographic phenomena, two fundamental and perennial concepts have maintained their significance. Theoretical and empirical understanding of context (places) and nonstationarity (spatial autocorrelation) continue to bedevil spatial analysis. Fundamental questions that relate to issues in defining and measuring "neighborhood", the social setting in which political interactions occur, remain only partly answered despite further work on these topics since the earlier version of this chapter appeared a decade ago. Relatedly, key uncertainties around the handling of spatial autocorrelation in geographic data are as important now as in early work within the revival of political geography.

Three developments in spatial analysis over the last decade have dramatically changed its nature and research outlook. Vast improvements in data availability, data analysis, and new methodological tools have made the technologies and techniques more accessible to a wider audience in geography and related social sciences. After reviewing the two primary tenets of spatial analysis and outlining the three recent changes that we observe, in this chapter we integrate these discussions into a stylized empirical example of understanding beliefs about violence in the North Caucasus region of Russia, a region which has seen continuous but localized violence since the collapse of the Soviet Union in the early 1990s. Our geographic information systems (GIS) analysis connects foundational concepts of political geography to the identification of novel methodological approaches and analytical tools using both individual- and aggregate-level data.

II. TIMELESS/CONSTANT FUNDAMENTAL CONCEPTS

a. Context and Place

Political geographers, despite their diversity of methodologies and topical foci, share the belief that the social settings of events influence individual and group level behaviors and attitudes. Political behaviors such as voting decisions, conflict about territorial control, political boundary delineation and demarcation, and public goods provision and allocation are a function of a constellation of influences that mix aggregate and individual factors at scales ranging from the individual to the locality to the national and international. Unlike other social sciences where individuals tend to be 'atomized' and interactions between them and their social settings unrecognized and unexamined, political geographers have highlighted the contextual effects of politics for over a century since Vidal de la Blache (1903) first documented and described the complex "milieux" (geographic settings) of France.

In this chapter, we illustrate possible extensions for classical approaches to identifying and measuring the effects of "place" in quantitative political geography, which we understand to be the "settings and scenes of everyday life" (Agnew, 1987, p. 5). A "structurationist" approach (the two-way interaction between individuals and their *milieux* that together lead to social outcomes) was the root of much foundational research pointing toward the importance of social context in human geography (e.g. Pred, 1983). Johnston and Pattie (1992) for example, showed that a “spatial polarization” of voting trends (the link between class and party support) had developed in the British electorate and that it pointed toward the importance of social context in shaping individual choices. Like many other studies of elections, Pattie and Johnston (2000) provided evidence that “people who talk together vote together”, and therefore, personal interactions in social settings condition political behaviors. Relatedly, in political science, Braybeck and Huckfeldt (2002) explicitly identified neighborhood contexts and measured the diffusion of information about candidates and issues in dispersed and condensed social networks that influence voting behavior. In a field experiment, Enos (2014) shows that increased inter- ethnic group contact in a baseline homogenous United States context can result in exclusionary social attitudes. While place-focused quantitative analysis has concentrated on elections or civic beliefs, it can of course be considered for other political phenomena such as violent conflict. “Why is it, for example, that political violence characterizes the political histories of some places, but not others? Often this may have been the product of place-specific repression, or the absence of other alternatives such as electoral politics” (Agnew, 1987, p. 60).

For most political geographers, understanding of place and context has moved away from what O’Loughlin (2000) decried as a geometric or ‘Cartesian’ view; this move now considers relational understandings of the settings in which people live their lives (Castree, 2004). Relative interactive measures of places, locales, and regions can be defined by links that are related to social identity, political economy, or migration. In this sense, place or context may be defined by group membership or network ties *in addition to* physical location (Massey, 1994). While relative and absolute concepts are often presented as mutually exclusive, and we acknowledge that our example below incorporates distance-based metrics, we believe that the concepts and their measures can simultaneously be accommodated in quantitative spatial analysis.

Though few other social scientists study similar topics, the focus of spatial analytical work in political geography is usually the elaboration of the nuanced effects of social forces expressed in localities at various spatial scales. This theme formed the centerpiece of the spatial analysis chapter in the previous edition of this book. The usual political science study tends to be deductively developed from a formal model (often a rational choice model) and oriented toward testing a single observable implication of a theory. Unfortunately, many non-geographers still rely on methods choices that “control away” the influence of contexts on human behavior, or, in Gould’s (1970) phrase, they are “throwing out the baby with the bathwater." For these researchers, a geographical analysis is a means to an end, and not the end itself. There is still a tendency in political science research to narrowly view space and place through proximity and/or contiguity lenses, essentially as geometric measures. Although there have been improvements in spatial analysis outside of geography (we highlight several examples below), this has not changed dramatically from the previous version of this *Companion* chapter ten years ago and the critique a decade before by the prominent political scientist, Gary King (King, 1996 in response to Agnew, 1996). Of course political geographers are not a-theoretical, but differing from political scientists, they approach hypothesis-testing with more emphasis on probing the empirics of data before predictive modeling. Disciplinary boundaries still sometimes represent silos of distinct conceptual frameworks for research.

The contextual theme is also commonly employed in other geographically-sensitive research fields. In studies of health outcomes, educational achievement, crime rates, and social mobility, it is common practice to directly assess the link between contextual qualities (such as neighborhood levels of poverty, racial or ethnic composition, political representation, etc.) and individual characteristics in nested data structures. For example, students learn within classrooms that are not identical because teachers are different, classrooms are in schools that may be different in terms of facilities and resources, and those schools are located in school districts that also vary dramatically due to tax base variations and other factors. A student's achievement partly depends upon his/her home location because observable traits like taxes earned from property values and dedicated to education vary widely in American metropolitan areas. Defining the boundary for capturing such contextual effects can influence results greatly as these varying spatial scales have qualities that can be measured and included in multi-level analysis.

Following standard econometric practices, most scholars in fact accept that geographic differences can be important and consequently, use statistical methods that incorporate a "fixed effects" approach. In this method, binary dummy variables are defined for each spatial unit, such as voting districts in a city, and thus can be numerous.

If the researcher is probing whether voters' incomes influence electoral participation, s/he would ideally have an accurate estimate for the effect of poverty after ensuring that other personal and contextual characteristics are not also influencing participation. The accessibility of polling stations in a constituency level could influence participation. If there are no data available for the number of polling station locations, comparing individuals statistically in one constituency only to individuals in that same constituency in a fixed effects model eliminates the possibility that differences between constituencies are also responsible for the changes in participation rates. The key problem with the dummy variable approach is that while the model controls for baseline differences between spatial units, the results of the analysis cannot tell us *how* the differences between units of observations matter.

Regardless of the chosen statistical model, the delineation of the areal dimensions of the spatial units matter greatly for the conclusions of spatial analysis. The Modifiable Areal Unit Problem (MAUP) is well-known but rarely examined methodically. MAUP was outlined by Gehlke and Biehl in 1934 but is most commonly associated with Openshaw (1983). The key MAUP issue is that the size and direction of a statistical association found at one scale of analysis (e.g. census tract) may not hold at others (e.g. county) (Openshaw and Taylor, 1979). Related to this inconsistency, changing the boundary shape of an area also changes relationships. For example, if we correlate socio-economic class and voting for the Republican party, the coefficient varies across hierarchical scales as a result of the number of data points and the geographic configuration of the districts, and therefore, we cannot be sure which coefficient is correct (Openshaw, 1996). MAUP dilemmas are different than, but related to, the more conceptual dilemma of the “ecological fallacy”. Researchers risk an ecological fallacy when they assume that aggregate-level data represent individual-level processes or phenomena (Robinson, 1950; Selvin, 1958). In essence, an ecological fallacy is MAUP at the finest resolution (individual level), and confounds explanations and outcomes across aggregate and individual scales.

As accurate measurement of context is central to spatial analysis in political geography and differing neighborhood bounds can influence results, more flexibility for researchers in defining the relevant neighborhood is warranted. Many empirically-derived metrics for selecting neighborhood dimensions exist (Root et al., 2009; Spielman et al., 2012) but fine-resolution data are required for these analytical tools to be effective. It is impossible to know if a larger unit of analysis is more appropriate than smaller units if one cannot examine and compare results at different levels. For analysts who view the world through the lens of spatial scales, decisions about the proper aggregations of information thus become very important. While it is possible to aggregate up from fine resolutions (location-based data) to coarse (large areal-unit data), disaggregating from coarse to fine resolutions is impossible (or at least very difficult, requiring calculations of uncertainty). If one wants to know how best to understand political violence in Afghanistan, knowing only the provincial-level violence rates limits one’s ability effectively to measure the local ebb and flow of conflict as it occurs relative to the border with Pakistan. We would also have no understanding of how violence is geographically distributed within ethno-regional enclaves at localized scales. With location level data, however, it is possible to aggregate points (and all of the attributes of individual violent events) to a zone defined by any absolute spatial reference (e.g. border) or relative spatial dimension (e.g. ethnic community dominance in a region) (O'Loughlin et al., 2010).

b. Nonstationary Data and Spatial Autocorrelation

In addition to considerations of context (*milieux*), another fundamental and timeless quality of spatial analysis in political geography is its nonstationarity quality. Data are nonstationary when the relationship between variables vary across locations within the dataset and are not consistent in different regions. As Tobler (1970) famously wrote, "everything is related to everything else, but near things are more related than distant things." Spatial dependencies (a term that reflects this distance effect) among units of analysis are important observable artifacts of largely unobserved social processes. Spatial analysts are interested in how, where and why data are geographically related. If human behavioral patterns cluster in space (and they almost always do), geographers strive to understand how and why, rather simply controlling for the trend as a nuisance that sullies an otherwise non-spatial model. Spatial terms in models attempt to capture these contextual effects.

Questions about spatially nonstationary data are answered inductively. The calculation of Local Indicators of Spatial Association (LISAs) has been the workhorse of spatial analysts for two decades and identify where a locally-specific measure of autocorrelation may not match a global or overall trend (Anselin, 1995). LISA statistics are based on a straightforward comparison between observed and simulated (expected uniform or random) spatial distributions. The Getis-Ord Gi* ("hot spot") statistic is one example of a LISA indicator (Ord and Getis, 1995). For a study of insurgency in Afghanistan between 2004 and 2009, using daily military reports of "significant activity" (SIGACTS), O'Loughlin and colleagues (2010), for small 25 km^2 gridcells, were able to identify changes in the locations and sizes of clusters of events initiated by both insurgent and coalition forces. As one would expect, some general overlap between the clusters can be observed and they also display differing locational trends. This is an example of using LISA-type statistics inductively, but such procedures can also be used to identify spatial trends that distort the results of non-spatial econometric analysis. Using a local clustering statistic for the residuals of a regression (difference between observed and predicted values), for example, can reveal clusters of unexplained variance in a dataset, suggesting that assumptions about the independence of observations (explained below) are violated.

Point pattern process identification, related to LISA indicators because they compare an observed spatial distribution to an expected (simulated) distribution, can be used to uncover local nuances and trends in political data (O'Loughlin, 2002). Recent advances in this style of analysis are extremely powerful, including the routines implemented in SaTScan, software initially designed to identify space-time clusters for epidemiological and other spatially-referenced health data (Kulldorff et al., 2005; Kulldorff and Information Management Services, 2009). The benefit of using SaTScan is that the temporal and spatial dimensions of the window that defines clustering are allowed to vary and the output is – at least conceptually – a cylinder with vertical and horizontal dimensions that can be mapped by location coordinates. For the study of violence in Afghanistan, this method delivers even more detail about the spatio-temporal dimensions of conflict than just a LISA indicator (O'Loughlin et al., 2010, p. 490).

With geographically-dependent data "knowing one value on the surface provides the observer with a better than random chance of predicting nearby values" (Gould, 1970, p. 444), thus violating the assumption that observation units are independently and identically distributed (IID). IID assumption violations can give researchers faulty assurances that the results of their quantitative analyses are valid because standard errors for testing the significance of a coefficient point estimate are artificially small (Anselin, 1988). After McCarty's (1954) analysis of geographic patterns of the vote for Wisconsin's right-wing senator Joseph McCarthy – one of the first studies explicitly promoting a spatial approach to statistical analysis – a steady stream of quantitative political geography methods textbooks and articles emerged that provide technical overviews of the confounding influences of spatial dependencies and the many possible solutions (Anselin, 2002; Anselin, 1988; Cliff and Ord, 1973).

One of the most straightforward treatments of spatial dependency is introducing an autoregressive (also called "spatial lag") term into a regression model. The classical estimator in regression analysis can be represented in simplistic form as $Y_i = \beta X_i + \varepsilon_i$, with, for observation *i*, *Y* represents the outcome variable, *X* the main independent variable, and ε the unobserved variation (error) in the relationship. The value of β quantifies the association between *X* and *Y*. More advanced models would include temporal dimensions and control variables that might also contribute to the outcome. A spatial autoregressive model is represented by $Y_i = \rho W_i Y + \beta X_i + \varepsilon_i$, with spatial weights matrix, W, used to define the presence of Y within a neighborhood surrounding observation *i.* W may be defined by proximity (distance) or contiguity (shared border). The estimate of ρ in the model can be interpreted in substantial terms. Interpreting the autoregressive term of the equation reveals a high level of predictability within a model (e.g. O'Loughlin et al., 2012).

In contrast to lag models that accept the role of geographic affinities in the data and tries to use them to understand human behavior, the "spatial error" solution to the problems posed by spatial dependency is an adjustment of the standard error of each coefficient estimate of the independent variables (β in the equation above) so that their influence on the outcome of interest is not overstated. In effect, a spatial error approach is similar to common adjustments to the error structure of a regression model (e.g. clustered errors at some spatial scale), but is specifically based on geographic distances or contiguities, defining the connections among the units (as is the case for W in the equation above). Because the spatial lag model introduces a new term into the model, consideration of the spatial contiguity effects has been explicitly incorporated into regression models of electoral choices for the past 35 years (O'Loughlin, 1981; O'Loughlin et al., 1994).

III. TRENDS/CHANGES IN SPATIAL ANALYSIS OVER THE LAST DECADE

a. Data Availability

While the foundational concepts remain perennially important, there have been major changes in the landscape of spatial analysis over the last decade. Primary among the changes helping to spur more spatial analysis is a dramatic increase in the availability of diverse types of data that are readily available for political geographic research. Following closely the publication of the earlier version of this chapter – on the cusp of important changes in data and analysis for human geography – Johnston et al. (2004, p. 367) wrote,

> "there are many hypotheses regarding neighborhood effects in the geographical and related literatures, but their successful testing has been hampered by the absence of data. In particular, analysts have lacked data on both individuals and their neighborhood milieux, which allow the interactions of different types of people in different types of local context to be explored."

In their article, Johnston et al. (2004) combined survey and electoral outcome data in "bespoke neighborhoods" in a manner that facilitated the creative identification of multiple scales at which social forces operate, thus effectively avoiding criticisms related to MAUP.

Quantitative political geography research had classically relied mostly on administrative unit polygon data, usually in electoral studies. However, quantitative analysis of political geographic subjects is increasingly going beyond only areal unit data to include a variety of formats and types (e.g. surveys in Linke 2013; Secor and O'Loughlin, 2005). Examples of the various formats include network (Radil and Flint, 2013), census small area (Verpoorten, 2012), point pattern location (Linke et al., 2012), satellite remote sensing (Henderson et al., 2012), roads (Zhukov, 2012), land use change detection (Witmer and O'Loughlin, 2009) and even mobile phone service data (Pierskalla and Hollenbach, 2013). Some research uses non-uniform polygons at a global scale to capture spatial demographic qualities of territories such as ethnic groupings (e.g. Wucherpfennig et al., 2011). Individual research teams most often collect these data, but they are increasingly available from the United Nations, international Non-Governmental Organizations and even national governmental agencies in the developing world.

The push toward sharing data among academics has also picked up pace alongside the increasingly mandated expectation of journals that data for quantitative analysis be made public for replication by other researchers in the field (King, 1995). The Dataverse Network Project at Harvard University (http://www.thedata.org), for example, is "free and open to all researchers worldwide to share, cite, reuse and archive research data." However, concerns regarding the confidential nature of individual data that are anchored to a specific location are growing at a time when social media, governmental, and private information collection is on the rise (VanWey et al., 2005).

A major "neogeography" shift in the availability of geospatial data has also seen volunteered private (eg. individual location) data used for a number of academic and nonacademic applications, in addition to the non-expert use of GIS to further the aims of a given community. The original rise of such data uses accompanied a revolution in the accessibility of web-based mapping applications, such as Google Maps' API and even the many extensions of the basic Google Maps (and Earth) interface(s). Where the non-traditional cartographic tools exist (i.e. those that do not rely on proprietary software like ESRI's ArcMap) and are accompanied by spatial data, the potential for "citizen science" emerges. Activist or otherwise community-based activities developed upon these platforms can have important real-world consequences. One example of such a technical neogeography within the realm of political spatial analysis would be the role of social media and cartographic mapping within the Syrian conflict. According to the *Washington Post*, for instance, activists and those sympathetic to the anti-Assad forces used Google extension Map Maker to rename the streets and locations in key Syrian cities to reflect opposition movement historical sympathies (Lynch, 2012). As a kind of hybrid academic-popular data collection and analysis effort, *Voix des Kivus* (directed by Peter van der Windt and Macarten Humphries at Columbia University) was a project that introduced truly innovative methods of data collection in the Democratic Republic of Congo (DRC). *Voix des Kivus* distributed mobile phones to communities across the war-torn eastern regions of DRC with the goal of capturing daily reports of violent incidents that occur below the radar of major private media, government, or non-governmental reporting. *Voix des Kivus* was based on a top-down organizational structure, but the principal operating premise is acceptance that everyday citizens play profoundly important roles in gathering the geographical data researchers in academic fields regularly use.

b. Data Use

A second recent trend in spatial analysis relates to the compilation and integration of multiple data formats. In traditional electoral geography, for example, analysis was often conducted with data available at a single scale (vote totals and predictors based on census data). In spatial analysis limited to a single scale of analysis and format of data important "place-influences" or contextual-level effects may not be captured effectively by distance and contiguity scores. To advance the measurement of possible local level influences upon outcomes of interest in political geography, hybrid units of analysis and mixed data structures are now more commonly used.

An analysis of conflict in Iraq can serve as an example of diverse data merging into a single GIS platform, allowing for the discovery of social relationships that are hidden when analysis is bound by a single source, analytical unit, or data format (Linke et al., 2012). Four different data formats and dimensions were merged - socioeconomic status of districts (survey data aggregated into vector polygons), ethnic group distributions (scanned paper map converted to vector polygons), violent event location data (latitude and longitude point coordinates), and satellite night-time lights to measure urbanization (raster image). Using a Granger causal effects estimator, a tit-for-tat dynamics of insurgent-regime forces reciprocity emerges where the actions of one side of the combat strongly predicted a timely and local response from the other. These associations varied across different spatio-temporal thresholds and across ethno-sectarian, income, and population density regions, a finding that emerged from the compilation of diverse data sources. A similar compilation of survey, violent events, and population data into a common sub-national unit of analysis for a conflict diffusion study of 16 countries in sub-Saharan Africa is found in Linke et al. (2014).

This kind of data integration on a specific spatial scale with complementary spatially-sensitive methods is developing rapidly. PRIO-GRID (Tollefson et al., 2012) is one example of a major effort to merge multiple freely available data sources for governmental, population, socioeconomic status, terrain, and ecological-climatological data into a ½ degree grid-cell unified structure for all world regions. While the basic dataset has a static spatial resolution (thus not automatically resolving MAUP issues), it represents a substantial advance in how conflict enquiry in political science is carried out with a move away from the "territorial trap" of the nation-state in favor of geographically-disaggregated research.

c. Analytical Tools

As well as the important changes in data availability and data use, recent years have also seen the dramatic rise in the number of statistical and graphical tools available for spatial analysts. Free platforms for statistical analysis supporting hundreds of procedures for spatial data management, mapping, and regression can now help researchers in innumerable ways after data are correctly formatted and organized (Bivand et al., 2008; Griffith and Paelinck, 2011; Plant, 2012). Many free stand alone software platforms exist for this kind of spatial analysis including GeoDa (Anselin et al., 2006) for Exploratory Spatial Data Analysis (ESDA) or LISA calculations (see also SatScan, noted above), as well as regression. Dedicated spatial analysis packages exist in the software platform *R* for managing classes of spatial data (**sp**), reading and writing shapefiles (**maptools**), creating and using spatial weights matrices (**spdep**), geostatistics and anisotropy (**gstat**), survey sampling points based on population distributions (**spsurvey**), and point-pattern processes (**spatial, spatstat**, **splancs**, and **spatialkernal**), multilevel modeling (**nlme4**), geographically weighed regression (**spgwr**), and advanced mapping and visualization (**ggplot2**). Researchers with coding experience can assemble a GIS dataset, execute their preferred statistical estimation, and display graphical results in a single program. This is a change from the past where software was mainly proprietary, and often GIS/mapping and statistical analysis were completed separately. While the learning curve for using command line programming interfaces in *R* can be steep, the payoffs are substantial in the long term.

IV. SPATIAL ANALYSIS OF VIOLENCE AND PUBLIC OPINION IN THE NORTH CAUCASUS OF RUSSIA

In this section we illustrate some of the principles and methods of spatial analysis in political geography research. Survey data for 2000 individuals from the North Caucasus of Russia were collected in December 2005. The survey was part of a comparative project with Bosnia-Herzegovina and was designed as a study of post-conflict attitudes toward group reconciliation and prospects for peaceful relations. The North Caucasus conflicts between 1994 and 1996, and restarting in 1999 were marked in the later years by guerrilla warfare, terrorist attacks, reprisals by the Russian forces, and a diffusion outward from its original core in Chechnya. Over time, the militancy took on a more Islamist character as its leaders declared a "shariat" (state run by Islamic law) for the Muslim republics across the region. By 2005, the conflict had waned significantly but violent events still occurred on a daily basis in Chechnya and neighboring republics. Our research question concerns the possible effects of the violence in the immediate area upon survey respondent attitudes.

The survey was designed as a geographically-stratified one and was conducted in 82 sampling points within the republics of North Ossetia, Dagestan, Karachevo-Cherkassia, Kabardino-Balkaria, and the territory of Stavropol’ (Figure 1) (Chechnya and Ingushetia were excluded from the study because high levels of violence at the time of the survey made it impossible.) We join the location of survey respondents with conflict data for the two years before the December 2005 survey. These violent events are coded from newspapers and the data include the exact location, day of event, the perpetrators and targets, and estimates of casualties. Coders reviewed thousands of Lexis-Nexis stories to identify events with enough information to ensure reliability in geographic precision. The violent event file is precisely georeferenced by UTM Zone X and Y coordinates after the violent event locations are identified, allowing for precise distance calculations in a projected coordinate system. The spatial nonstationarity of the violent events data has been illustrated by kernel density surfaces and conditional probabilities of reciprocal violence in O’Loughlin and Witmer (2012).

a. Aggregating the Data Formats

With the goal of measuring multiple settings or neighborhoods of violence, we allow the spatial boundaries around survey sampling points to vary in making violent event aggregations. Using a range of dimensions around a location allows us to address the uncertainty of the relevant context for a respondent and illustrate the MAUP choices. On the map in Figure 1, we identify five sampling locations (Oktyabrskoe, Prokhladny, Zavodskoi, Buinaksk, and Cherkessk) that we profile in the tabled event counts of neighborhood violence measures below (see Figure 2). We select each of these because they have experienced varying levels of violence prior to the survey (some extremely violent and other comparatively peaceful). Additionally, the five locations represent each of the republics within the broader study area. We choose these only for the purposes of illustrating the data aggregation steps and the variation of violence that is captured across space-time dimensions. For conclusions based upon the models below, the selection of these five illustrative towns has no meaningful consequence.

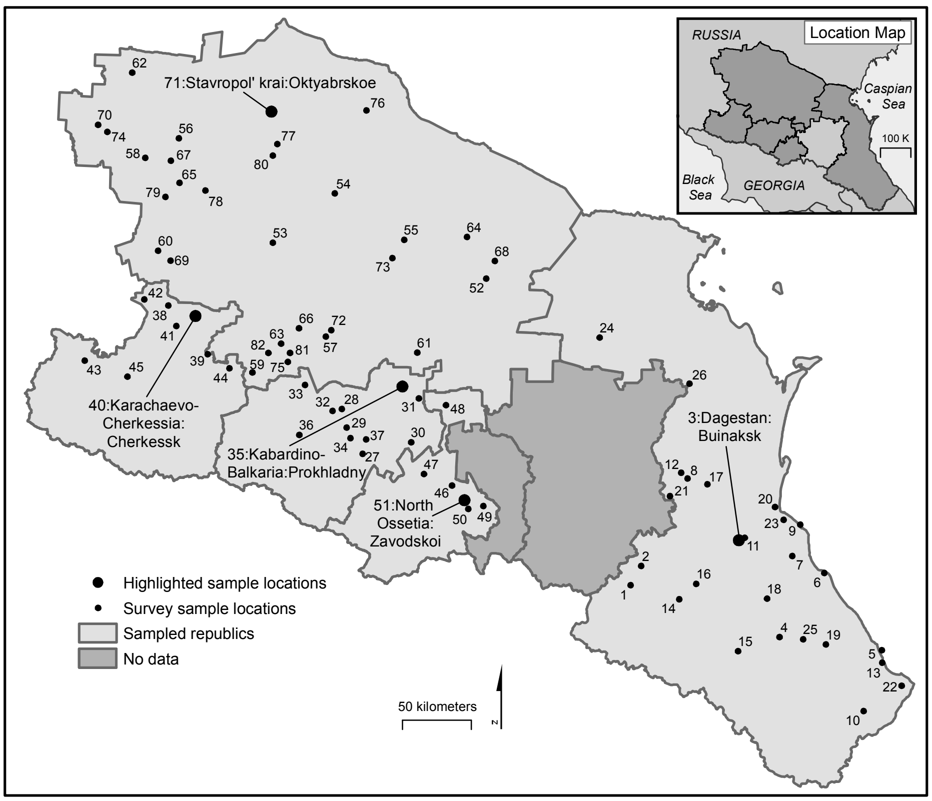

Figure 1: The 82 primary sampling locations for the 2005 survey in five republics of the North Caucasus. To ease comparison, the sample sites identified in the figure are also highlighted on the graphic in Figure 2.

For this illustration, we use eight three-month time slices up to a maximum of two years and ten spatial thresholds (10km – 100km). Specifically, we first measure the distance between each of the 82 survey locations and every violent event location (total locations of violent incidents number 1,367). We use the *pointDistance* function in *R* and count the number of events that has taken place within each time-space buffer around those locations. Each row in Figure 2 represents a survey sample point, whose numbers 1-82 correspond to the locations on the map. The thicker black horizontal lines represent borders between republics. The shade of each cell represents the conflict event count (logged because of some high values) of each space-time buffer. This heatmap presentation of quantitative data can be very helpful in exploratory spatial data analysis. Illustrating a large number of dimensions (here 82 space X 80 time cells = 6560 cells) as a kind of choropleth matrix is more graphically arresting and helpful than standard tabular results.

The profiles of all locations show the expected trends. As the spatial buffer becomes very large (100km) the number of events increases. Many locations have no violence recorded nearby when the 10km distance is used, but at 100km, almost every location has some violence. Secondarily, as the temporal range used to define our pairing of violence data with locations expands from three months to 24 months, a clear trend toward higher conflict event counts is also visible.

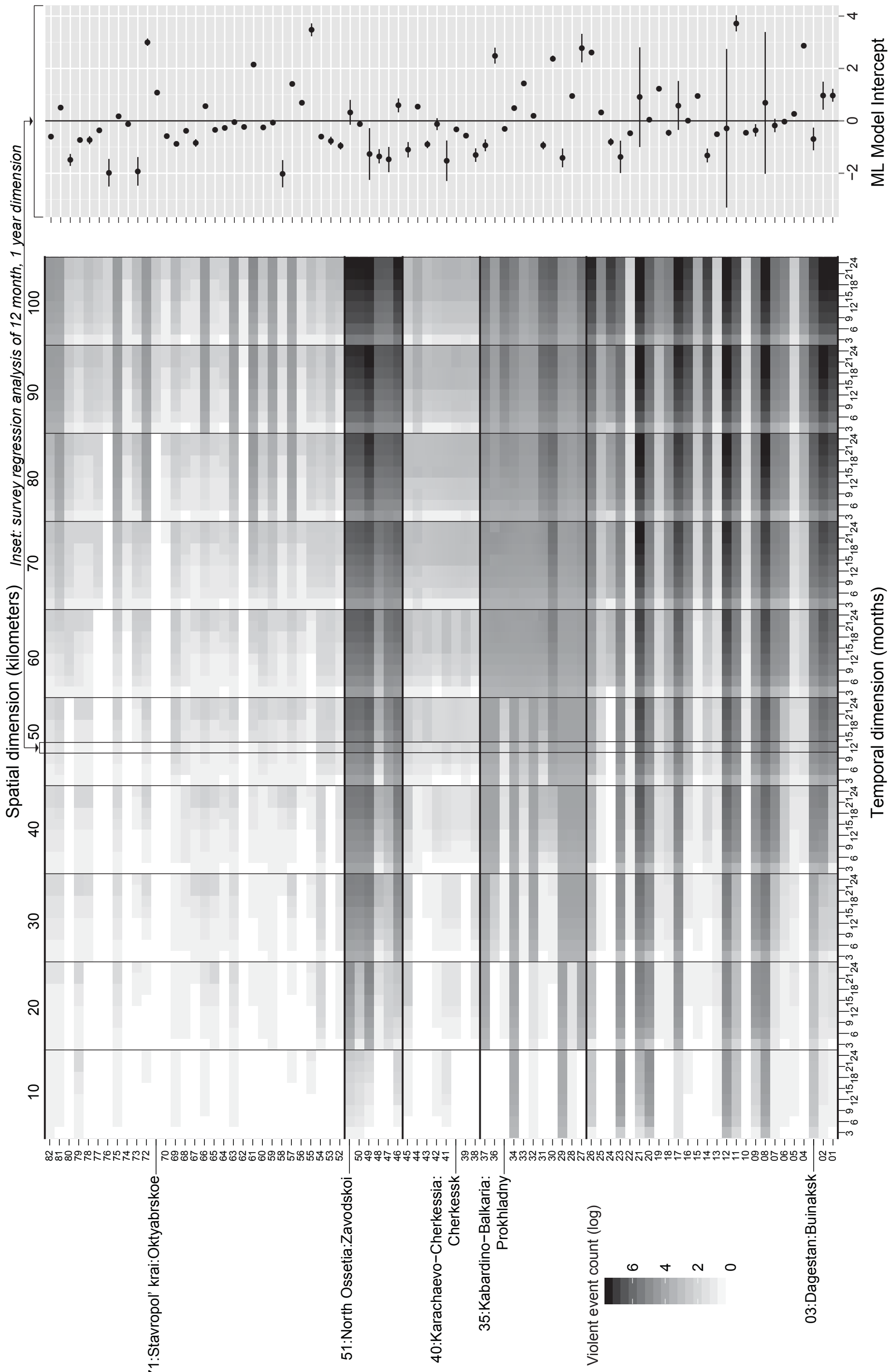

Figure 2: The number of violent events aggregated within 80 spatio-temporal dimensions (3 months breaks by 10 km boundaries) for each of 82 survey sample locations in the North Caucasus. Survey locations (horizontal axis) correspond to the map locations in Figure 1. Thicker black horizontal lines represent borders of the republics. For legibility, not all labels are shown. The smaller graphical figure (right) is an illustration of the multilevel/hierarchical regression model intercept for a single space-time dimension (here, 12 months and 50km). Similar variation exists for the quantitative modeling of survey responses in each column representing a space-time dimension.

Five illustrative locations demonstrate different overall trends in violence. Zavodskoi near Ingushetia (number 51) is quite violent, with conflict taking place across nearly all temporal and spatial dimensions (and with very high rates at coarse boundary definitions). In dramatic contrast, Oktyabroskoe (76) in northern Stavropol' territory is far from the core of violence in Chechnya and thus, relatively peaceful. Cherkessk (40) represents a medium level of violence with no conflict found at very fine temporal and spatial resolutions, but accumulates as distances increase. From this graphical display, we make the straightforward conclusion that the "violence neighborhood" surrounding locations varies dramatically between localities and according to threshold delimitation. The more important consideration relates to our earlier discussion of MAUP and it remains unclear what is the correct range for delimiting the violence context.

b. Varying Intercepts in Statistical Models

The multilevel modeling approach is one that allows the evaluation of individual and contextual effects to be measured and compared since it is now almost axiomatic in political geography that "context matters" in influencing individual choices and behaviors. Multilevel modeling has not achieved prominence in political geography compared to other social sciences, especially epidemiology, criminology, and other public policy research. By allowing the basic relationship between the outcome of interest and a key predictor to vary across settings/contexts – that are known to be qualitatively different in unobserved, unmeasurable, and unknown ways – we allow for the influence of contextual factors to be incorporated into the analysis of survey respondent attitudes as potentially important place-influences.

For our illustration, we analyze answers to the question: "what of the following listed is the most serious danger facing the peoples of the North Caucasus in the next 5 years?" More than crime, ethnic separatism, corruption, and lack of economic development and unemployment, a plurality of people (42%) claimed that "terrorist actions and military conflicts" were the biggest problem. We code those individuals as one and other potential threats as zero for a binary response in a logistic regression model. To take account for alternative explanations of the fear of terrorism, we control for individual-level factors (income and employment status, education, age, and gender).

The relationship we model can be represented as $Y_i = \alpha_0 + \beta \mathrm{X}_i + Z_i + \alpha_s + \ \varepsilon_i$, for observations $i = \{1, 2 \ldots 2000\}$ and $\alpha_s \sim N(0, \sigma_\alpha^2)$, for survey location level $s = \{1, 2 \ldots 82\}$ where α_0 is a random effect, Y_i is the outcome, β is the neighborhood violence indicator, and X_i is the vector of individual level control variables that could confound the association (β) between X_i and Y_i. Unexplained variance is captured in ε_i, the error term. A non-hierarchical generalized linear model (GLM; for the binary outcome) model would exclude the α_s term in its empirical estimate, the component of the model that allows the statistical intercept of the relationship under investigation to vary across units at some designated hierarchical scale (survey locations, in our case). In a random effects model, the variance of the intercepts can be informative. Heterogeneity at finer resolutions is likely to be more substantial than at coarse spatial resolutions.

To indicate if the relationship between individual concerns about terrorism/violence and level of violence in the locality changes by survey sampling point, a plot of the intercepts of the regression model for each of the second level units is helpful. (A plot for the third republic level would be warranted in a more comprehensive research article). In the right-hand side of Figure 2, we present the statistical association between the variable capturing rates of local violence and a respondent's worry about terrorism. As an example, we selected a spatio-temporal dimension (12 months and 50 kms) near the middle of the time and distance ranges. Each location identified in the map (Figure 1) and cloropleth matrix (Figure 2, left) is included in this graphical plot.

A key component of our analysis is identifying whether the variation between sampling point intercepts is statistically significant. The intercept values vary from roughly -2 to more than +3.5 indicating a wide divergence in the relationship between violence and respondent worries about terrorism across the study region. The places where a respondent lives, in other words, has a major bearing on what a respondent ranks as an important worry, even after controlling for individual level influences such as age or education. To see if the random effects estimator (Table 1) is statistically significant, ANOVA (analysis of variance) of the log-likelihood for different models is appropriate.

c. Illustrating MLM Results

Careful consideration of geographical context and scale can change substantial interpretations of statistical analysis. For the results in Table 1 below, we use the **lme4** package and the generalized linear model (for binary outcomes) function *glmer*. Using a logistical regression functional form, a statistically significant influence of contextual violence on worries about terrorism emerges (columns b-d). For every additional violent event that took place within 100km of the survey location in the preceding year, the odds of respondent worry about terrorism increase marginally. Since the number of events within the space-time buffer ranges from 2-1319, 0-454, and 0-135 for 100km, 50km, and 20km ranges, respectively, this is a noteworthy effect even though the coefficient estimate is small. Drawing conclusions from the basic non-hierarchical model, one would conclude that a context of violence and insecurity influences worries about terrorism and that this relationship is robust across definitions of spatial neighborhood. However, our multilevel model approach shows that the basic relationship between regional violence and worries about terrorism varies widely across survey points as indicated by the intercepts plotted in Figure 2 above.

Influence of violence taking place at varying spatio-temporal scales, including random effects estimators (models e -g)

	a) controls only			b) 20km			c) 50km			d) 100km			e) RE 20km			f) RE 50km			g) RE 100km		
	Est	StdEr	OR	Est	StdEr	OR	Est	StdEr	OR	Est	StdEr	OR	Est	StdEr	OR	Est	StdEr	OR	Est	StdEr	OR
(Intercept)	- .29	.10 *	0.748	- .37	.10 *	0.688	- .38	.10 *	0.685	- .39	.10 *	0.677	- .30	.19	0.738	- .32	.19 *	0.729	- .40	.19 *	0.670
12mo 20km total				**.00**	**.00 ***	**1.003**							**.00**	**.00**	**1.002**						
12mo 50km total							**.00**	**.00 ***	**1.002**							**.00**	**.00**	**1.001**			
12mo 100km total										**.00**	**.00 ***	**1.001**							**.00**	**.00 ***	**1.001**
Education	- .09	.10	0.917	- .09	.10	0.918	- .09	.10	0.911	- .10	.10	0.908	- .15	.11	0.859	- .15	.11	0.860	- .15	.11	0.861
Male	- .10	.09	0.906	- .10	.09	0.904	- .10	.09	0.903	- .10	.09	0.903	- .12	.10	0.885	- .12	.10	0.886	- .12	.10	0.885
Poor	.31	.10 *	1.357	.32	.10 *	1.371	.29	.10 *	1.334	.28	.10 *	1.328	.12	.11	1.130	.12	.11	1.132	.12	.11	1.129
Rural	- .04	.09	0.961	.00	.09	0.997	- .02	.09	0.980	- .04	.09	0.959	.03	.20	1.034	.00	.21	0.997	.00	.20	1.004
Tense interview	- .20	.10 *	0.819	- .21	.10 *	0.808	- .22	.10 *	0.804	- .22	.10 *	0.802	- .20	.11 *	0.819	- .20	.11 *	0.816	- .21	.11 *	0.809
AIC	2719.51			2715.52			2711.64			2707.29			2620.86			2622.95			2621.78		
RE significant	N/A			N/A			N/A			N/A			Yes			Yes			Yes		

RE = Random intercept model; N/A = not applicable; * = Statistically significant at 10%

Table 1: The influence of violence taking place in various space-time scales upon survey respondent concerns about terrorism and war in his/her area. Statistical estimator is a random intercept multilevel binary outcome logistical regression model, with intercepts varying by the 82 survey locations. Bold table values are highlight as the main estimates of interest (effect of violence neighborhood upon attitudes).

The multilevel model results for place-based statistical intercept variations tell a modified but important story (columns e-g of Table 1). In contrast to our initial conclusion, the link between regional violence and fear of terrorism is statistically significant only at the 100km threshold. In other words, a small spatial range (20kms) does not yield significant results because the number of events at such a local level is small. It is the broader, more regional concentration of violence that influences respondent perceptions. In political geography, such place-specific influences are important and consistently appear in different locales and settings, and we should investigate whether our models remain consistent and robust after taking such effects into account. While our example is specific to the violent North Caucasus in 2005 and the distance effect would be different in other settings, we cannot know *a priori* what the relevant thresholds are until we measure a variety of them and show the variation in the model of contextual effects.

V. CONCLUSION

Political geography is a changing and dynamic discipline, as it should be. However, certain thematic continuities are threaded from the post World War II resurgence of to the present. Spatial analysis in human (including political) geography must wholeheartedly embrace a nuanced approach to understanding contexts, places and social settings of human behavior. It must also fully recognize the spatial dependencies measured by different metrics between locales and administrative units, including the linkages among individuals living in those places. While we reiterate the importance of these perennial and timeless issues for political geography – we have referred to early recognition of contextual effects in 1903 and quantitative work from 1934 – there have also been remarkable changes over the past decade in the availability and use of data for spatial social science. The empowering changes can foster robust and valuable improvements to quantitative research but must be accompanied with thoughtful and meticulous training in the assumptions that underlie statistical methods. A firm understanding of theory and concepts in political geography can effectively be merged with cautious and careful quantitative approaches to our discipline; the two are not mutually exclusive and hybrid mixed methods research agendas will hopefully continue to illustrate this combination in the future.

References:

Agnew, J. 1987. *Place and Politics*. London: Allen and Unwin.

Agnew, J. 1996. Mapping politics: How context counts in electoral geography. *Political Geography, 15, 129-146.*

Anselin, L. 1988. *Spatial Econometrics: Methods and Models.* Dordrecht: Kluwer.

Anselin, L. 2002. Under the hood. Issues in the specification and interpretation of spatial regression models. *Agricultural Economics*, 27, 247-267.

Anselin, L., Syabri, I., and Kho, Y. 2006. GeoDa: An introduction to spatial data analysis. *Geographical Analysis*, 38, 5-22.

Braybeck, B. and Huckfeldt, R. 2002. Urban contexts, spatially dispersed networks, and the diffusion of political information. *Political Geography*, 21, 195-220.

Bivand, R. S., Pebesma, E. J., and Gomez-Rubio, V. 2008. *Applied Spatial Data Analysis in R*. New York, NY: Springer.

Castree, N. 2004. Differential geographies: Place, indigenous rights and 'local' resources. *Political Geography*, 22, 133-167.

Enos, R. D. 2014. The causal effect of prolonged intergroup contact on exclusionary attitudes: A test using public transportation in homogenous communities. *Proceedings of the National Academy of Sciences,* forthcoming, doi: 10.1073/pnas.1317670111.

Gehlke, C. E. and Biehl, K. 1934. Certain effects of grouping upon the size of the correlation coefficient census tract material. *Journal of the American Statistical Association*, 29, 169-170.

Getis, A. and Ord, J. K. 1992. The analysis of spatial association by use of distance statistics. *Geographical Analysis,* 24, 189-206.

Gould, P. 1970. Is statistix inferens the geographical name for a wild goose? *Economic Geography*, 46, 439 – 448.

Griffith, D. A. and Paelinck, J. H. P. 2011. *Non-standard Spatial Statistics and Spatial Econometrics*. New York, NY: Springer.

Henderson, J. V., Storeygard, A., and Weil, D. N. 2012. Measuring economic growth from outer space. *American Economic Review*, 102, 994-1028.

Johnston, R., Jones, K., Burgess, S., Propper, C., Sarker, R., and Bolster, A. 2004. Scale, factor analyses, and neighborhood effects. *Geographical Analysis,* 36, 350 – 368.

King, G. 1995. Replication, replication. *PS: Political Science and Politics*, 28, 443-499.

King, G. 1996. Why context should not count. *Political Geography*, 15, 159-164.

Kulldorff, M., Heffernan, R., Hartman, J., Assuncao, R., and Mostashari, F. 2005. A space-time permutation scan statistic for disease outbreak detection. *PLoS Medicine*, 2, 216-224.

Kulldorff, M. and Information Management Services. 2009. *SaTScan: Software for the Spatial and Space-Time Scan Statistics, v 9.0.1,* 2009 [http://www.satscan.org/].

Linke, A. M. 2013. The aftermath of an election crisis: Kenyan attitudes and the influences of individual-level and locality violence. *Political Geography*, 37, 5-17.

Linke, A. M., Witmer, F., and O'Loughlin, J. 2012. Space-time Granger Analysis of the war in Iraq: A study of coalition and insurgent action-reaction. *International Interactions*, 38, 402-425.

Linke, A. M., Schutte, S., and Buhaug, H. 2014. Population attitudes and the spread of Political violence in Africa. Paper presented at the annual International Studies Association meeting, Toronto, CA.

Lynch, Colum. 2012. "Syrian opposition seeks to wipe the Assad name off of the map – via Google." *The Washington Post*, 14 February.

Massey, D. 1994. *Space, Place, and Gender*. Minneapolis: University of Minnesota Press.

O'Loughlin, J. 1981. The neighborhood effect in urban voting surfaces: An international comparison. In A. D. Burnett and P. J. Taylor (eds.) *Political Studies from Spatial Perspectives: Anglo-American Essays in Political Geography*. Chichester, UK: John Wiley and Sons, 357-388.

O'Loughlin, J. 2000. Geography as space and geography as place: the divide between political science and political geography continues. *Geopolitics*, 5, 126-137.

O'Loughlin, J. 2002. The electoral geography of Weimar Germany: Exploratory Spatial Data Analysis (ESDA) of Protestant Support for the Nazi Party. *Political Analysis*, 10, 217-243.

O'Loughlin, J., Flint, C., and Anselin, L. 1994. The geography of the Nazi vote: Context, confession and class in the Reichstag election of 1930. *Annals of the Association of American Geographers*, 84, 351-380.

O'Loughlin, J., Witmer, F., Linke, A. M., and Thorwardson, N. 2010. Peering into the fog of war: the geography of the Wikileaks Afghanistan war logs 2004-2009. *Eurasian Geography and Economics*, 51, 437-71.

O'Loughlin, J. and Witmer, F. 2011. The localized geographies of violence in the North Caucasus of Russia, 1999-2007. *Annals of the Association of American Geographers*, 101, 178-201.

O'Loughlin, J. and Witmer, F. 2012. The diffusion of conflicts in the North Caucasus of Russia, 1999-2010. *Environment and Planning A*, 44, 2379-2396.

O'Louglin, J., Witmer, F., Linke, A. M., Laing, A., Gettleman, A., and Dudhia, J. 2012. Climate variability and conflict risk in East Africa, 1990-2009. *Proceedings of the National Academy of Sciences*, 109, 18344-18394.

Openshaw, S. 1983. *The Modifiable Areal Unit Problem*. Norfolk, Geo Books.

Openshaw, S. 1996. Developing GIS-relevant zone-based spatial analysis methods. In P. Longley and M. Batty (eds.) *Spatial Analysis: Modeling in a GIS Environment*. New York: Wiley, 55-73.

Openshaw, S. and Taylor, P. 1979. A million or so correlation coefficients: three experiments on the modifiable areal unit problem. In Wrigley, N. (ed.) *Statistical Applications in the Spatial Sciences*. London: Pion, 127-144.

Ord, J. K., and Getis, A. 1995. Local spatial autocorrelation statistics: distributional issues and application. *Geographical Analysis,* 27, 286-306.

Pattie, C. and Johnston, R. 2000. 'People who talk together vote together': An exploration of contextual effects in Great Britain. *Annals of the Association of American Geographers*, 90, 41-66.

Pierskalla, J. and Hollenbach, F. 2013. Technology and collective action: The effect of cell phone coverage on political violence in Africa. *American Political Science Review*, May 2013, doi:10.1017/S0003055413000075.

Plant, R. E. 2012. *Spatial Data Analysis in Ecology and Agriculture Using R*. Boca Raton, FL: CRC Press.

Pred, A. 1983. Structuration and place: on the becoming of sense and place and structure of feeling. *Journal for the Theory of Social Behavior,* 13, 45-68.

Radil, S. and Flint, C. 2013. Exiles and arms: the territorial practices of state making and war diffusion in post-Cold War Africa. *Territory, Politics, Governance*, 1, 183-202.

Robinson, W. S. 1950. Ecological correlations and the behavior of individuals. *American Sociological Review*. 15, 351-357.

Root, E. D., Meyer, R., and Emch, M. 2009. Evidence of localized clustering of gastroschisis in North Carolina, 1999-2004. *Social Science and Medicine*, 68, 1361-1367.

Secor, A. J. and O'Loughlin, J. 2005. Social and political trust in Istanbul and Moscow: A comparison of individual and neighborhood effects. *Transactions, Institute of British Geographers*, 30, 66-82.

Selvin, H. C. 1958. Durkheim's suicide and problems of empirical research. *American Journal of Sociology,* 63, 607-619.

Spielman, S. E., Yoo, E-H., and Linkletter, C. 2013. "Neighborhood contexts, health, and behavior: understanding the role of scale and residential sorting." *Environment and Planning B*, 40, 489-506.

Tobler, W. 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46, 234-240.

Tollefson, A. F., Strand, H., and Buhaug, H. 2012. PRIO-GRID: A unified spatial data structure. *Journal of Peace Research*, 49, 363-374.

VanWey, L. K., Rindfuss, R. R., Gutmann, M. P., Entwisle, B., and Balk, D. L. 2005. Confidentiality and spatially explicit data: Concerns and challenges. *Proceedings of the National Academy of Sciences*, 102, 15337-15342.

Vidal de La Blanche, P. 1903. *Tableau de la geographie de la France.* Paris, Hachette (reprinted: Paris, Tallandier, 1979).

Verpoorten, M. 2012. Detecting hidden violence: The spatial distribution of excess mortality in Rwanda. *Political Geography*, 31, 44-56.

Witmer, F. and O'Loughlin, J. 2009. Satellite Data Methods and Application in the Evaluation of War Outcomes: Abandoned Agricultural land in Bosnia-Herzegovina after the 1992-1995 conflict. *Annals of the Association of American Geographers,* 99, 1033-1044.

Wucherpfennig, J., Weidmann, N. B, Girardin, L., Cederman, L-E., and Wimmer, A. 2011. Politically relevant ethnic groups across space and time: Introducing the geoEPR dataset. *Conflict Management and Peace Science*, 28, 423-437.

Zhukov, Y. M. 2012. Roads and the diffusion of insurgent violence: the logistics of conflict in Russia's North Caucasus. *Political Geography*, 31, 144-156.