Civil Wars: Prospects and Problems with the use of Local Indicators

Jan Ketil Rød[1] and Halvard Buhaug[2]

[1] Department of Geography, Norwegian University of Science and Technology (NTNU) and Centre for the Study of Civil War (CSCW), International Peace Research Institute, Oslo (PRIO)
jan.rod@svt.ntnu.no
[2] Centre for the Study of Civil War (CSCW),
International Peace Research Institute, Oslo (PRIO)
halvardb@prio.no

Abstract
Civil wars rarely span throughout the territories of the host countries. More frequently, they confine to specific sub-national regions characterized by the present of certain geographic factors favorable for warfare. In order to avoid the ecological fallacy, or explaining local phenomena with country-level measures, an increasing number of quantitative studies of civil war are applying disaggregated data and research designs. Although such approaches are certainly promising, they introduce or emphasize problems related to classical statistical inference. A fundamental underlying assumption is that of independence, that the units of analysis are unrelated to each other. With geographical data, this assumption is questionable, and particularly so for disaggregated observations. Disaggregating also presents another challenge, known as the modifiable areal unit problem (MAUP). In this paper, we investigate possible effects of spatial autocorrelation and MAUP by conducting a series of disaggregated analyses on the incidence of civil war in Africa, 1970–2004. To test for the zoning effect of MAUP, we use two grids with identical resolution (100 x 100 km) but with different zoning (52% overlap) to generate alternative datasets, which are then analyzed and evaluated. The scale effect of MAUP and the influence of spatial autocorrelation are explored by estimating similar regression models on five alternative samples, generated from grids with a resolution of 50, 100 (twofold), 150, and 200 km, respectively.

1. Introduction

Studies on interstate war have shown that conflicts are spatially dependent and that they form clusters in space [25, 4]. O'Loughlin claims that 'nations located in proximity to each other are more likely to go to war than nations located far apart' [19] and concluded that 'spatial factors are as important as military expenditures and are more important than commonly used political and economic predictors in explaining war behaviour' [19]. These studies show that geographic location does appear to play an important role in interstate war behaviour and we believe the role of geographic location is even more important in civil wars than in interstate wars. This is because civil wars rarely span throughout the territories of the host countries. More frequently, they confine to specific sub-national regions characterized by the present of certain geographic factors favorable for warfare. A location may be favourable for a rebel group if the place provide refuge from governmental control (e.g. remoteness from state capital, nearness to international boundary, rough terrain), or rich on mineral (such as diamonds) and thus attractive areas to control for combatants who need to finance their warfare. If the motive for fighting is to overthrow the ruling government, the fighting will take place where the power is situated, mostly in or near the capital. These are just a few examples on how location plays a role in the study of civil war.

Still, few empirical studies have successfully established a relation between civil war and the geographic distribution of relevant physical and social factors. Is this because general patterns do not exist and that, consequently, techniques designed to discover such patterns are inappropriate? We believe that stable patterns and relationships may in fact exist, but that in order to recognize them, we need to concentrate on variables drawn from the same level of analysis as the internal conflicts themselves. This is an approach several scholars currently adopt, namely to situate the analysis at a level corresponding with their hypothesis. Based on a literature review of empirical literature on civil war, Buhaug and Lujala [6] find that most studies ignore local conditions and instead use country aggregates in their analysis. Empirical studies on civil armed conflicts have traditionally used country level variables such as level of democracy, population density and measures of national poverty. Nation-level poverty is often measured by gross national product per capita and is recognised as one of the most prominent and robust factor associated with the occurrence of civil war [9, 10, 11, 12 and 16]. There is, however, disagreement on how to interpret these results. Some maintain that gross national product per capita is a proxy for state strength, meaning that richer regimes are better able to monitor the population and conduct effective counterinsurgencies [12]. Another popular inter-pretation is that in poor countries there are more young men with low wages who are willing to join a rebel group when other better paid work is difficult to find [10].

However, from observing that most armed conflicts take place in poor countries to draw the conclusion that lower-income persons are more likely to join a rebel group, would be an ecological fallacy, ignoring that a strong aggregated relationship does not necessarily explain lower level phenomena. The ecological fallacy is a problem that has been known for a long time. A classical study by Gehlke and Biehl [13] noted that correlation coefficients tend to increase with the level of geographic aggregation when census data are analyzed. Although there is assembled evidence in epidemiology and elsewhere suggesting that the correlation coefficient is more seriously affected than regression parameters [15], O'Sullivan and Unwin [20] note that regression relationships are strengthen by aggregation. Rogerson add that 'what is significant at one spatial scale may not be significant at another' [22]. Even though the effects of ecological fallacy may still not be particular well understood, compared to studies using country statistics, we may expect that the results from regression analysis will show a less strong relationship when using measures at disaggregated levels than if we continue to use country measures.

It is therefore necessary to stress that the observation that most conflicts occur in poor countries should not lead to conclusions that insurgencies start in the poorest areas or that those who fight are the poorest. The "opportunity cost" proposition might still be correct, but its conclusion should, ideally, be based on individual level data. Unfortunately, we do not have reliable individual level data on income for our spatio-temporal domain. If such data indeed were collected for the African continent (for instance as an additional variable in the demographic health survey (DHS)) it would probably be available with uncontrollably large imprecision. We are therefore unable to test the opportunity cost proposition in this paper.

While the test of the opportunity cost proposition should be based on data on an individual level, other research questions force the use of data aggregated to the country level. As argued by Buhaug and Rød [7] statistical studies that aim to explore the government side in civil war and focus on political aspects of the regime will necessarily have to be conducted at the country/government level. Such studies would probably be based on characteristics that are constant within countries such as the number of years since the last regime change, type of political system, and whether or not the country is a major oil exporter. Empirical studies that aim to explore the side of the insurgents will, on the other hand, profit from disaggregated data as their hypotheses often pertain to sub-national conditions, for instance, that rebels who seek refuge in mountainous terrain, closed forests, near international borders and far from governmental control (the capital) are better able to withstand a military superior opposition. Consequently, rebel groups will take advantage of such terrain, whenever available. Buhaug and Rød's [7] as well as Buhaug and Lujala [6] therefore argue that whenever we investigate theories of civil war that have an element of geography, we should seriously consider abandoning the habitual country level of analysis in favour of a disaggregated approach. Otherwise, we are likely to fall prey to the ecological fallacy by explaining local phenomena from aggregated data. In order to avoid the ecological fallacy and to test whether the theoretical local mechanism in fact also holds as empirical local facts, increasingly number of authors now tends to disaggregate their variables. Examples of disaggregated studies on civil armed conflicts also include Raleigh and Urdal [21], and Theisen and Bransegg [23] which all collect disaggregated variables onto 100 x 100 km squares, Hegre and Raleigh [16] who use a finer resolution 8.6 x 8.6 km, and Østby et al, [27] who use administrative regions.

We believe disaggregated approaches to be promising, but there are some pitfalls when conventional statistics are employed on geographical data. The most important of these have the headings modifiable area unit problem (MAUP) and spatial autocorrelation. We would be even more convinced about the excellence of a disaggregated research design if we would know how such an analysis would be influenced if based on units which had been zoned differently. In this paper we replicate Buhaug and Rød's [7] main findings, but estimate the logistic regression parameters based on units with both varying scale (which determine the size of each spatial unit) and partition (planar offset of the spatial units given the scale of analysis).

2. Replicating Disaggregated Study Using Various Zoning Schemes

Buhaug and Rød [7] placed their analysis on the onset of civil armed conflicts on a sub-national scale, using squared polygons with 100x100 km resolution. They based the sample of conflicts on an improved version of the Uppsala/PRIO dataset, where the conflicts are represented by polygons to reflect the geographical area of the battle zones. Squares that overlap with a conflict polygon are coded as having a civil war onset in the initial year of the conflict. Similarly, they assigned figures to the squares by overlaying spatial layers with features often associated with civil war; like for instance, population density, amount of rough terrain, and distances to international boundary, capital and various natural resources. Further, Buhaug and Rød [7] distinguished between conflicts over territory (secession) and conflicts over governance (coups, revolutions) since these types of conflicts are likely to be shaped partly by different conditions. Their findings offer support to the notion that governmental and territorial armed conflicts are shaped by different causal mechanisms[1] as their main findings were:

- territorial conflicts are much more likely to occur in sparsely populated regions near international borders and far from the capital
- governmental conflicts occur predominantly in densely populated areas near the capital

These findings motivate us, in this paper, to focus essentially on the three most important explanatory variables from their study: population counts, distance to capital and distance to international border. As Buhaug and Rød [7], we also separate the response variable (civil armed conflicts) into territorial and governmental, but as this paper reflects work in progress we are mainly reporting findings from our analysis on territorial armed conflicts. For the moment, we look at onset of territorial armed conflicts only, on the African continent extended to the period from 1970 to 2004. Any square that overlaps with a conflict zone is, in a dummy variable, coded as having a conflict in the given year. Each unit is further assigned specific values on the three space-varying and potentially conflict-promoting variables: population counts and relative location (distance to capital and distance to international boundary). The population data are taken from UNEP-GRID,[2] who has generated a raster representation of population counts for Africa at a resolution of 2.5 arc minutes (approximately 5 km at equator). The data are available for every decade since 1960 and give estimates for the number of inhabitants in each grid square. We applied linear interpolation to fill in data for missing years and linear extrapolation to fill in data for the years after 2000. As usual, we take the natural logarithm of the measure (modified to population in 1,000s) to reduce outlier bias. We measured distances from the centroids at each square to the capital in the country and to the nearest point at the international border. Islands, such as Madagascar and the Comoros, are arbitrary assigned a distance to international border equal to 1.000 km. Since transportation cost has not changed dramatically since 1970 in this region, we do not use any

---

[1] See also Buhaug [5] for additional empirical support for this distinction.

[2] United Nations Environmental Program – Global Resource Information Database, data available from: http://grid2.cr.usgs.gov/datasets/datalist.php3#unep.

estimate exponent for various years.[3] Both of the distance measures are calculated along geodesic curves rather than Euclidean distances. Since equidistant map projection allows you to generate accurate distance measures only between specific points, geodesic distances are really the only way to get consistently accurate distances measured when the distance goes over significant portions of the earth. For large countries such as Sudan, the Democratic Republic of Congo (DRC) or Algeria, the difference between a geodesic and a Euclidean distance may be 10 % or more. As with the population variable, we take the natural logarithm of the distance measures to account for an expected non-linear effect.

There were several boundary changes and changes of capital during the spatial-temporal domain of our study, due to for instance the independence of Eritrea and Namibia. We have accounted for these changes by generating country-year variables. We then collapsed the time-series data (1970–2004) to a single cross section, generating median values for all explanatory variables, mean values for the spatial lags, and maximum values for the dichotomous conflict indicators. The dependent variables in the collapsed datasets thus separate between square cells without conflict since 1970 and cells that have hosted conflict in one or more years in the analysis period.

3. Elaborating on the Pitfalls with Disaggregated Studies

As with Buhaug and Rød [7], we use tessellated squares as our unit of analysis. There are at least three difficulties involved with the use of disaggregated tessellated units: the problem of selecting appropriate form, size and location for the spatial units. Although there are several alternative geometric forms possible, quadrate shapes are commonly used in empirical disciplines involving fieldwork like geography and biology. We adopted the convention on using regular squares as the form for our spatial units. The size of the squares (often referred to as scale or resolution) should fit the positional accuracy of our dataset. The gridded population counts we used have a spatial resolution at about 5 km, the positional accuracy of the polygons representing areas affected by armed civil conflict are rather coarse, we would estimate an accuracy approximately at 50 km which we selected as the lowest resolution for the squares. However, if the spatial unit system in a particular study was specified differently (e.g. 100 km or 200 km resolution), we might observe different relationships, a problem referred to as the modifiable areal unit problem, MAUP [26].

In order to examine the sensitivity of our results to modifiable areal units we divide the spatial domain of our analysis – Africa – into squares with a resolution of 50x50 km, 100x100 km, 150x150 km and 200x200 km. As indicated in Figure 1, each square polygon is assigned to one country only. Squares that cover international boundaries and thus overlap several countries are defined as belonging to the country

[3] O'Loughlin [19] experienced difficulties in specifying theoretically the value of a distance exponent in a study on international conflicts between 1816 and 1980, using an absolute distance measure would be suspect since the costs of overcoming distance have changed dramatically over the 164 year long study period.

that dominates the area of the square polygon. As there are several small countries in Africa we carefully located the 200 km square to assure that each country became represented with at least one square polygon, such as Rwanda and Burundi (see Figure 1).

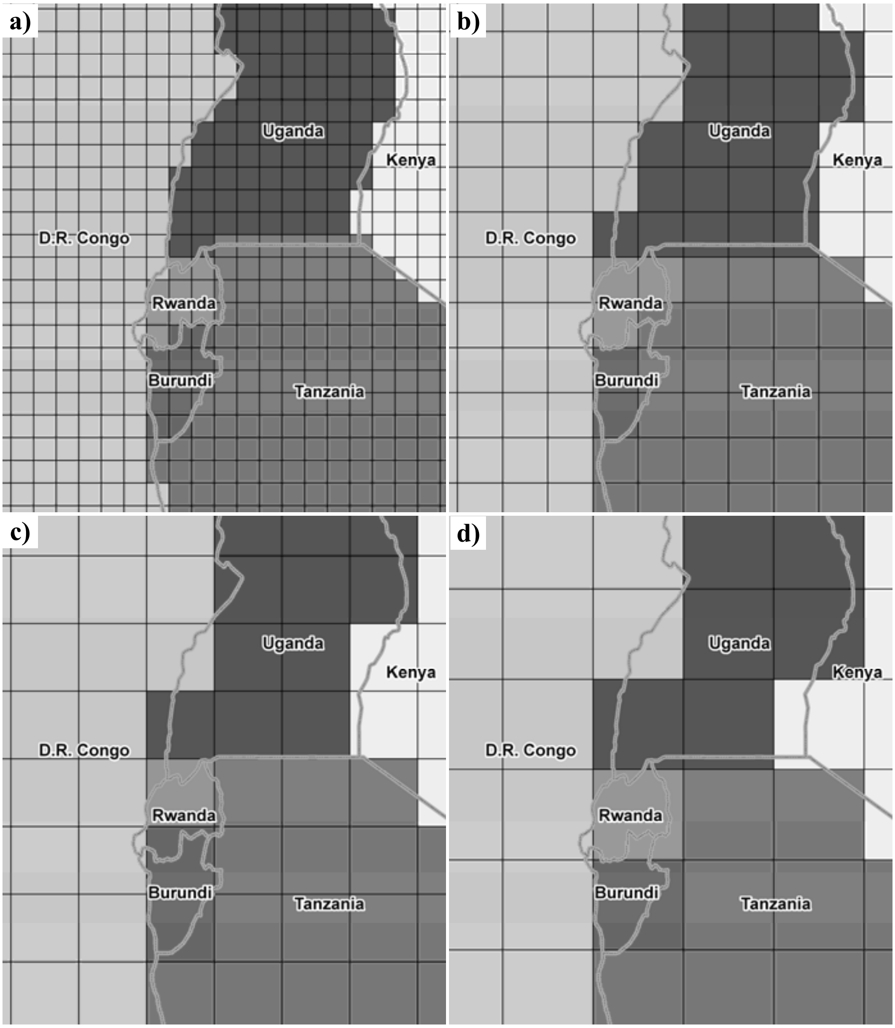

Fig. 1. Each square polygon is assigned to one country. Examples from the four resolutions: a) 50 km, b) 100 km, c) 150 km, and d) 200 km

Since there are several capitals being close to each other[4] it was impossible to locate the 200 km fishnet without ending up with squares containing two capitals.

[4] Kinshasa and Brazzaville is about 20 km apart, Porto Novo and Lagos about 80 km apart.

Four capitals were therefore slightly relocated[5] for the 200 km fishnet to make the calculation of distances from capitals feasible. One 100 km fishnet and the 50 km fishnet were aligned perfectly with the 200 km fishnet. For the squares with 100 km resolution, we have two alternative representations with 52 % overlap (offset of 42 km W and 11 km S) with 3225 and 3218 units respectively.

3.1 Modifiable Areal Unit Problem and Spatial Autocorrelation

The term 'modifiable' is used because neither the choice of number of spatial units (the scale of the analysis) nor their particular configuration (the selected partitioning or zoning given the scale of analysis) is fundamental and any one of a number of other choices could have been made [15]. In general therefore, the MAUP consist of two different aspects that should be appreciated. The first aspect is related to geographic scale whereas the latter refer to the placement of zonal boundaries. When data at different scales or resolution bring about inconsistent results, it is referred to as the scale effect. When the study area is tessellated into units of different spatial configuration schemes from which data variables are compiled, it is referred to as the zoning effect. When we replace the 100 km grid with grids of larger cells, the results of our analysis are likely to be different. In order to be able to investigate the scaling effect, we have generated variables at four different resolutions (50 km, 100 km, 150 km and 200 km).

Spatial autocorrelation is a term referring to the fact that data from locations near one another in space are more likely to be similar than data from locations remote from each other [20]. Geographers often call this Tobler's 'first law of geography', that 'everything is related to everything else, but near things are more related than distant things' [24]. The prefix 'auto' means self and autocorrelation can thus somewhat loosely be defined as self-correlation. As such it involves a single variable, and refers to the correlation between pairs of observation made on this single variable. This type of correlation arises because realizations of the variable in question are ordered in some way [14]. Unfortunately, using conventional statistics on spatial data one of the assumptions of the classical regression model is violated; that data consist of independent observation without any such ordering. That means for our dataset on armed conflicts that whether a particular square polygon is coded as affected by conflict is not independent on how the neighbouring squares are coded.

There are several methods that can be used to detect spatial autocorrelation. The global Moran's I statistics [18] is a classical as well as a very common method used to measure the degree of spatial autocorrelation in area data. Both dichotomous and ratio level data can be examined for significant spatial autocorrelation using Moran's index. We use it to examine spatial autocorrelation first on the binary dependent variables (onset of territorial/governmental armed conflicts which are coded "1" (conflict) or "0" (no conflict)), and then on the residuals resulting from the logistic regression which are at a ratio level. Moran's index is calculated as follows:

[5] Conakry (Guinea) 42 km W, Porto Novo (Benin) 18 km N, Malabo (Eq Guinea) 17 km SW and Maseru (Lesotho) 2 km W.

$$I = \frac{n\sum_{i}^{n}\sum_{j}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{(\sum_{i}^{n}\sum_{j}^{n} w_{ij})\sum(y_i - \bar{y})^2} \tag{1}$$

where there are n regions (or square cells in our case), w_{ij} (the weights) is a measure of the spatial proximity between square i and j. Often, the weighting function w_{ij} is a measure of binary connectivity, where $w_{ij} = 1$ if i and j are adjacent and $w_{ij} = 0$ otherwise, or the weighting function is an inverse distance measure ($1/d_{ij}$). The values of Moran's index ranges from -1 to +1 and is interpreted similarly as a correlation coefficient. A Moran's index value near +1.0 indicates clustering; an index value near -1.0 indicates dispersion. Using the python script available in ArcGIS 9.2, we calculated Moran's *I* using weight matrix files based on binary rook connectivity.

The weight matrix files constrain the number of neighbours to be considered and we have made additional restrictions on the rook connectivity when calculating Moran's index. By definition, civil armed conflicts take place within the boundaries of a state. Our GIS representation of the geographical extensions of armed conflicts does therefore ends at international borders and the weight matrix files reflect this situation since a square is coded as being a neighbour to another square only if it is within the same country[6]. The entries for these neighbour relations are not included in the weight matrix files[7] and switch off the corresponding covariance term and thus excluding it in the determination of the index. We have not (yet) found anything in the literature indicating that this is violating the use of Moran's *I*, and we have neither found anything describing similar situations. There is much literature on the use of local indicators of spatial associations (LISA) and there is also a local measure of Moran's *I* developed [2] which is adopted for situations where there is a local clustering around one or few 'hot spots'. At this stage of the research, we have not investigated whether these measures would be applicable to our dataset, but we plan to do so in the near future.

We also row standardized the weights so that row sums equal 1 as a result of dividing each entry in a row by the sum of row values (e.g. number of cell the entry cell is rook-connected to). Table 1 below reports the results of the calculated Moran's index for the dependent variables onset of territorial and governmental civil armed conflicts.

The python scripts also return expected values and z-scores for Moran's index. The expected value of Moran's *I* is the value that would be obtained if there were no spatial autocorrelation. The expected value is defined by:

---

[6] During our sample period (1970–2004) the outline of the international boundaries saw major changes involving Namibia, South Africa, Eritrea and Ethiopia. This does not influence the regression analysis, but do affect the calculation of spatial autocorrelation based on the collapsed dataset. In the collapsed dataset, a country code is assigned each square polygon representing the country which the cell belonged to for the most part of the period.

[7] Commonly these would be entered as zeros in the weight matrix file. However, in the python script available in ArcGIS 9.2 only neighbour entries codes "1" is needed when determining Moran's index.

$$E(I) = -\frac{1}{n-1} \tag{2}$$

Since all our cases have large 'n', the statistical significance of any departure from the expected value can be tested using z-scores, defined as:

$$Z = \frac{I - E(I)}{\sigma} \tag{3}$$

The denominator in formula 3, σ, is the standard deviation of I. The computation of the standard deviation of I is based on the null hypothesis (zero spatial autocorrelation) of randomization, which is generally used in geographic studies. The randomization hypothesis state that the probability that the observed values are arranged in a random manner given all possible arrangements.

Table 1. Spatial dependencies in the dependent variables

Zoning	n	E(I)	I_{terr}	Z_{terr}	I_{gov}	Z_{gov}
50	12473	-0.00008	0.933	138,850	0.936	139.210
100	3225	-0.00031	0.942	67.722	0.911	65.501
100 b	3218	-0.00031	0.938	67.320	0.905	64.939
150	1480	-0.00068	0.915	42.385	0.871	40.305
200	862	-0.00116	0,908	30.884	0.824	28.024

The result from Table 1 shows robust and very high values for Moran's index at all zoning schemes as well as some increase in value with an increased resolution (particularly for the variable representing governmental armed conflicts). The third column in Table 1 shows the expected values of Moran's I. The z scores are very high for all resolutions assessing that the observed clustering is statistically significant and thus not random.

Spatial Lagged Regression

To counter the problem of spatial autocorrelated dependent variables, Anselin [1] suggests two alternative models, a spatial lag model in which the dependent variable exhibits positive spatial autocorrelation and a spatial error model in which the errors in the regression are spatially autocorrelated. The first approach is often also called spatial autoregressive model and is the approach we follow in this paper. The *spatial lag* (or autoregressive element) is an extra variable expressing the neighbour relationship. Our spatial lagged binary logistical regression model can be expressed as:

$$y_i = a_i + b_1 x_1 + b_2 x_2 + \ldots b_n x_n + b_{n+1} \sum_{j=1}^{m} w_{ij} y_j + \varepsilon_i \tag{4}$$

where y_i is the conflict onset dummy of square polygon *i*; a_i is the intercept value, b_1 - b_n are the regression coefficients for the *n* variables; and b_{n+1} is the autoregressive coefficient for the term $\sum w_{ij} y_j$ (summed from $j = 1$ to *m*), the influence of civil armed conflict in the *m* neighbouring square polygons *j* on square polygon *i* through the use of the w_{ij} contiguity matrix described earlier.

We have used three different spatial lags with increased detail regarding contiguity:

- Conflict in country (whether at least one other square in the same country is affected by conflict)
- Queen contiguity, dummy (whether at least one of the eight 1. order neighbour cell's are affected by conflict)
- Weighted spatial lag based on first, second and third orders contiguity

The two first spatial lags are easily constructed. Spatial lags involving higher order neighbours can be constructed by special software for spatial statistical analysis such as GeoDa [3], but are unable to construct spatial lags where neighbouring cells at opposite sides of an international border are not regarded as ontiguous. The principle for this blocking of contiguous is best explained by reference to Figure 2.

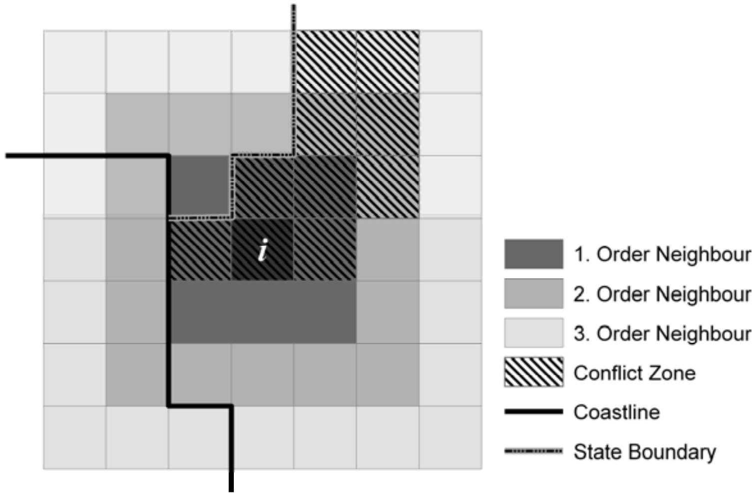

Fig. 2. Principle on the design of the spatial weighted lag which is based on first, second and third order neighbours. Neighbour across state boundary are excluded

Figure 2 shows cell *i* with its first, second and third order neighbours. A conflict zone is represented by the diagonal line pattern. The solid black line is a coastline and the grey line with black dotting is a state border. Only grid cell neighbours which lay within the same country and at the mainland are investigated. The grid cell *i* in Figure 2 has therefore seven first order neighbours (four of them in conflict), nine second order neighbours (three of them in conflict), and twelve third order neighbours (two of them in conflict). Each cell's neighbour relations (NBR) are then calculated using the following formula:

$$NBR = w_1 \frac{nc_1}{n_1} + w_2 \frac{nc_2}{n_2} + w_3 \frac{nc_3}{n_3} \tag{5}$$

where n_1, n_2 and n_3 are the number of first, second and third order neighbours respectively, nc_1, nc_2 and nc_3 are the number of first, second and third order neighbours in conflicts, and w_1, w_2, and w_3 are weights used to reduce the influence by distance. We used $w_1 = 0.6$, $w_2 = 0.3$, and $w_3 = 0.1$.

We express the neighbour relationship for cell *i* in Figure 2 as follows:

$$NBR = 0.6\frac{4}{7} + 0.3\frac{3}{9} + 0.1\frac{2}{12} \approx 0.46 \tag{6}$$

This value is calculated for each cell in each conflict year for both governmental and territorial armed conflicts for the five sets of zoning representations: a typical job for a computer. It was done by using batch processing python scripts in ArcGIS.

Results

We used Stata 9.2 to design and run 15 spatially lagged logistical regression models in which the dependent variable exhibits positive spatial autocorrelation; one model using each of the three spatial lags and one model for each of the five zoning systems. The differences between the coefficients resulting from the various models were very small. For territorial armed conflicts, the estimates for distance to capital are positive for all models suggesting that the risk of territorial civil armed conflicts are higher in regions far away from the capital. Distance to capital was the only variable that turned out to be significant and its coefficients for the 15 models are shown in Figure 3. It was significant at 1% level using the country lag through all five resolutions, but was significant only at the 5% level and only for the 150 km zoning using queen's lag. Finally, for the models using the weighted lag, distance to capital was significant at the 5% level for all of the zoning systems except the 50 km fishnet. These results do not support the notion that regression relationships are strengthen by aggregation, but they do give some support to the notion that what is significant at one spatial scale may not be significant at another.

One assumption of regression analysis as applied to spatial data is that the residuals are independent and thus are not spatially autocorrelated; there should be no spatial pattern to the errors. We therefore tested for spatial autocorrelation for the residuals resulting from each of the models. For the models having onset of territorial armed conflict as dependent variable, the result is shown in Figure 4.

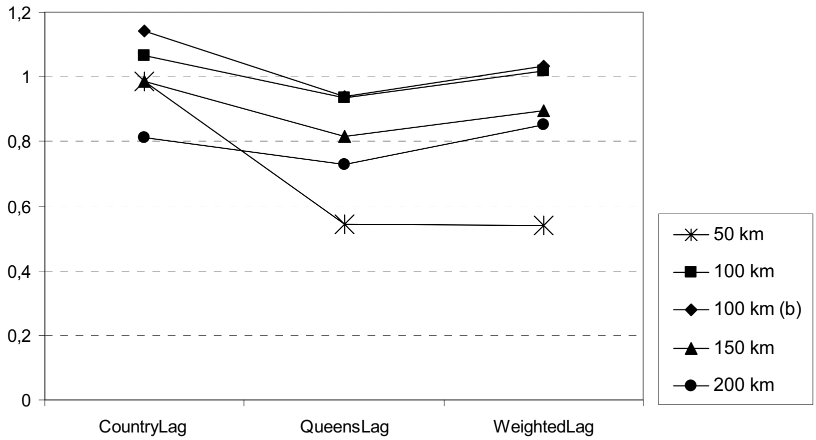

Fig. 3. Regression coefficients for the variable "Distance to Capital

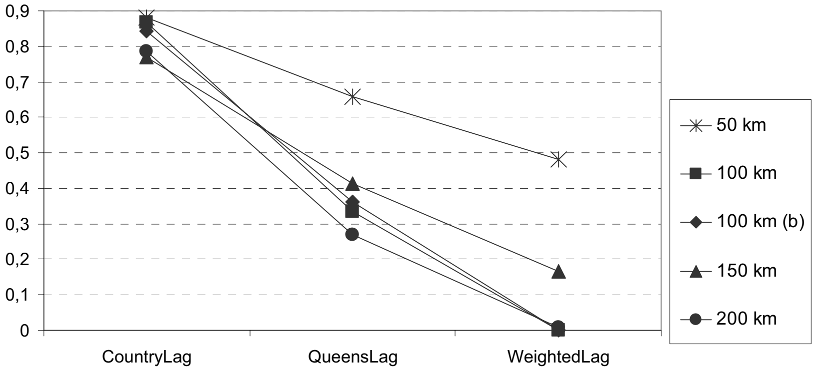

Fig. 4. Spatial dependencies in the residuals resulting from the 15 autoregressive models on the onset of territorial armed conflicts

As for the dependent variables (see Table 1), the residuals were tested for spatial dependencies using row standardized contiguity matrix files where country boundaries block neighbourhoods. Figure 4 demonstrates a general trend of decreasing Moran's *I* values for models where the spatial lags have an increased detail regarding neighbour relationships. For all models except three, the residuals of our regression equation (equation 4) are all positively and significantly autocorrelated, thereby violating one of the assumptions of regression. Such positive autocorrelation is the result of the misspecification of the form of the relationship, the absence of a significant predictor, or the absence of a geographic element [8, 19]. As the onset of territorial civil armed conflict is highly complex and has an unpredictable nature, we do not consider our sample of explanatory variables to be complete. Parts of the considerable amount of spatial autocorrelation in the residuals may therefore be due to misspecification of the models. According to O'Loughlin [19], the addition of a spatial autoregressive component frequently solves the problem of autocorrelation in the error terms. As seen from Figure 4, only three of the models have an index values close to zero (no spatial dependency) and with z-scores indicating that the residuals are random. These three models are all using the weighted lag: the two models with 100 km resolution and the 200 km resolution. Still, it is premature to interpret the absence of significant trend in the residuals for these models to mean that the models have been properly defined.

Concluding Remarks

Contrary to our expectations, neither the scaling effect nor the zoning effect of MAUP influences the regression results on the onset of territorial armed conflicts. The differences between the effects of the explanatory variables are small through the various zoning schemes (it remains to be seen if similar results are obtained from models on the onset of governmental armed conflicts). However, according to our results one should not underestimate the effect of spatial autocorrelation. The three different spatial autoregressive components we have used demonstrate a striking difference in their ability to solve the problem of autocorrelation in the error terms. We would suggest that care should be taken in designing contiguity indices for the lagged variable and that several measures of neighbour relationships are preferable to a single one. One assumption of regression analysis as applied to spatial data is that the residuals are independent and thus are not spatially autocorrelated the test for spatial autocorrelation should be used to judge the "goodness" of a model.

References

1. Anselin, L. 1988. *Spatial econometrics: Methods and models.* Dordrecht, Kluwer.
2. Anselin, L. 1995. Local Indicators of Spatial Association - LISA. *Geographical Analysis* 27(2): 93-115.
3. Anselin, L., I. Syabri and Y. Kho. 2006. GeoDa: An Introduction to Spatial Data Analysis. *Geographical Analysis* 38(1): 5-22.
4. Braithwaite, A. 2005. Location, Location, Location... Identifying Hot Spots of International Conflict. *International Interactions* 31(3): 251-273.
5. Buhaug, H. 2006. Relative Capability and Rebel Objective in Civil War. *Journal of Peace Research* 43(6): 691 - 708.
6. Buhaug, H. and P. Lujala. 2005. Accounting for scale: Measuring geography in quantitative studies of civil war. *Political Geography* 24(4): 399-418.
7. Buhaug, H. and J. K. Rød. 2006. Local Determinants of African Civil Wars, 1970 - 2001. *Political Geography* 25(3): 315-335.
8. Cliff, A. D. and J. K. Ord. 1981. *Spatial processes: models & application*. London, Pion.
9. Collier, P. and A. Hoeffler. 2002. On the Incidence of Civil War in Africa. *Journal of Conflict Resolution* 46(1): 13-28.

10. Collier, P. and A. Hoeffler. 2004. Greed and grievance in civil war. *Oxford Economic Papers* 56: 563-595.
11. de Soysa, I. 2000. The Resource Curse: Are Civil Wars Driven by Rapacity or Paucity? In. *Greed and Grievance. Economic Agendas in Civil Wars*. Boulder, Lynne Rienner Publishers**:** 113-135.
12. Fearon, J. D. and D. D. Laitin. 2003. Ethnicity, Insurgency, and Civil War. *American Political Science Review* 97(1): 75-90.
13. Gehlke, C. E. and K. Biehl. 1934. Certain effects of grouping upon the size of the correlation coefficient in census tract material. *Journal of the American Statistical Association.* 29(185): 169-170.
14. Griffith, D. A. 1987. *Spatial Autocorrelation. A Primer*, Association of American Geographers.
15. Haining, R. 2003. *Spatial Data Analysis. Theory and Practice.* Cambridge, Cambridge University Press.
16. Hegre, H. and C. Raleigh. 2005. Population Size, Concentration, and Civil War. A Geographically Disaggregated Analysis. Summer Meeting of the Polarization and Conflict Project. Konstanz.
17. Hegre, H., T. Ellingsen, S. Gates and N. P. Gleditsch. 2001. Toward a Democratic Civil Peace? Democracy, Political Change, and Civil War, 1816 - 1992. *American Political Science Review* 95(1): 33-48.
18. Moran, P. A. P. 1948. The Interpretation of Statistical Maps. *Journal of the Royal Statistical Society. B.* 10(2): 243-251.
19. O'Loughlin, J. 1986. Spatial Models of International Conflicts: Extending Current Theories of War Behavior. *Annals of the Association of American Geographers* 76(1): 63-80.
20. O'Sullivan, D. and D. J. Unwin. 2003. *Geographic Information Analysis*. Hoboken, Wiley.
21. Raleigh, C. and H. Urdal. 2006. Climate Change, Environmental Degradation and Armed Conflict. 47th Annual Convention of the International Studies Association. San Diego, CA.
22. Rogerson, P. A. 2006. *Statistical Methods for Geography.* London, Sage.
23. Theisen, O. M. and K. B. Brandsegg. 2007. The Environment and Non-State Conflicts in Sub-Saharan Africa. 48th Annual Convention of the International Studies Association. Chicago.
24. Tobler, W. R. 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography* 46: 234 - 240.
25. Ward, M. and K. S. Gleditsch. 2002. Location, Location, Location: An MCMC Approach to Modeling the Spatial Context of War and Peace. *Political Analysis* 10(2): 244-260.
26. Wrigley, N., T. Holt, D. Steel and M. Tranmer. 1996. Analysing, modelling, and resolving the ecological fallacy. In: P. Longley and M. Batty. *Spatial Analysis: Modelling in a GIS environment.* Cambridge, Geoinformation International**:** 23-40.
27. Østby, G., R. Nordås and J. K. Rød. 2006. Regional Inequalities and Civil Conflict in 21 Sub-Saharan African Countries, 1986–2004. Polarization and Conflict. Nicosia, Cyprus.