SPATIAL-TEMPORAL MODELING OF CIVIL WAR: THE EXAMPLE OF BOSNIA

NILS B. WEIDMANN AND MICHAEL D. WARD

Abstract. We describe a spatially and temporally autoregressive discrete regression model, detail and implement an algorithm for estimating the parameters of such a model, following the framework of Geyer and Thompson (1992), as recently updated by Zhu, Zheng, Carroll and Aukema (2008); Zhu, Rasmussen, Møller, Aukema and Raffa (2008); Zheng and Zhu (2008). This model is applied to geo-located data on attributes and conflict events in Bosnia over the period from March 1992 through October 1995. We present an $\mathcal{R}$ program library to estimate this class of models. Results show that there is a strong spatial as well as temporal dimension to the outbreak of civil conflicts in Bosnia, dynamics in space and time that standard OLS-type implementations completely miss. Using this approach it is no longer necessary to assume that either the spatial or the temporal dependencies in conflict data are exogenous in order to create predictive and inferential models of civil conflicts. Substantively, we show via inference, simulation, and animations that conflict events in Bosnia did diffuse spatially as well as temporally.

1. The Spatial-temporal Model

Discrete spatial temporal data characterize most models of conflict, both at the international and domestic levels. The gold-standard approach in political science appears to focus on the temporal dependencies in such data, at the expense of the spatial dependence. In the fields of applied statistics, there are two general stepping off points for developing models of such data. One approach follows the model based geo-statistics approach found in Diggle, Moyeed and Tawn (1998) and uses a form of generalized linear mixed models to examine spatial dependencies among geo-located events. The spatial-temporal component is modeled by a Gaussian process that captures the dependencies via an autocorrelation function that embodies both spatial and temporal lags.

A second approach follows from the breakthrough of Besag (1972, 1974) in using a lattice based framework for a Markov Random Field, which models the discrete outcome at one location conditional on the observed outcomes at neighboring locations. This approach–known as autologistic regression–was widely employed for decades by using a pseudo-likelihood approach (i.e., a normal logistic regression). The autologistic approach is fairly straightforward. We detail it formally herein. Let $\mathbf{Y}_t = (Y_{1,t} \dots Y_{N,t})'$ represent whether there is a conflict on the observed spatial arrangement of units at time t; $Y_{i,t}$ is the existence of conflict at location i at time t.

Date: September 15, 2008, version 0.9, conference draft, do not cite.

Authors' Note: This was prepared for the GROW-Net Conference on "Disaggregating Civil War: Transitions, Governance and Intervention," held in Zurich, 18-20 September, 2008. We would like to thank Jun Zhu for sharing code from his STARM package with us. Jan Ketil Rød has our gratitude for his help in obtaining a digital representation of the Bosnian municipalities.

By definition, $Y_{i,t} \in \{0,1\}$, 1 corresponding to the presence of conflict. Covariates are given in $X_{k,i,t}$. If the dependent variable is governed by a Markov random field, then conditional on the structure of the neighborhood $\mathcal{N}$,

$$Pr(Y_{i,t} \mid Y_{j,t} : j \neq i; j \in \mathcal{N}_i; \mathbf{Y}_{t'=t-1 \ldots t-S})$$

Since we have a binary, discrete response, a logistic formulation with a conditional formulation that is governed by the Bernoulli distribution is used to model the conditional probability of conflict, $p_{i,t}$:

$$\text{logit}(p_{i,t}) = \sum_{k=0}^{K} \theta_k X_{k,i,t} + \frac{1}{2} \sum_{l=1}^{L} \theta_{K+l} \sum_{j \in \mathcal{N}_i} Y_{j,t} + \sum_{s=1}^{S} \theta_{K+L+s} Y_{i,t-s}$$

where θ_K are the coefficients for the covariates, θ_{K+L} are the spatial autoregressive coefficients, and the remaining θ_{K+L+S} correspond to the temporal autoregressive components. By the Hammersley-Clifford Theorem[1] the joint distribution can be inferred from:

$$\begin{aligned} Pr(Y_t \mid Y_{t'} : t' = t-1 \ldots t-S) = {} & c(\mathbf{Y}_{t-1} \ldots \mathbf{Y}_{t-S}; \theta)^{-1} \times \\ & \exp\Bigg\{ \sum_{i=1}^{I} \sum_{k=0}^{K} \theta_k X_{k,i,t} Y_{i,t} + \\ & \frac{1}{2} \sum_{i=1}^{I} \sum_{l=1}^{L} \theta_{K+l} \sum_{j \in \mathcal{N}_i} Y_{i,t} Y_{j,t} + \\ & \sum_{i=1}^{I} \sum_{s=1}^{S} \theta_{K+L+s} Y_{i,t} Y_{t,t-s} \Bigg\} \end{aligned}$$

Unfortunately, $c(\mathbf{Y}_{t-1} \ldots \mathbf{Y}_{t-S}; \theta)^{-1}$ is a normalizing constant that doesn't have a closed form, and it is impossible for any real problem to evaluate it dynamically through an iterative estimation procedure such as MCMC or MLE. This problem is typically resolved by using a strategy suggested by Geyer and Thompson (1992) and used by Hoeting, Leecaster and Bowden (1999), Huffer and Wu (1998), Gumpertz, Wu and Pye (2000), and Ward and Gleditsch (2002), among others. This approach approximates a likelihood ratio as follows. Let $\mathbf{Z}_t$ represent the expansion of all the covariates, allowing a simpler statement of the log-likelihood function:

$$\mathcal{L}(\theta) = - \sum_{t=S+1}^{T} \log c(\mathbf{Y}_{t-1} \ldots \mathbf{Y}_{t-s}; \theta) + \sum_{t=S+1}^{T} \theta' \mathbf{Z}_t$$

Consider a parameter vector $\psi = (\psi_0 \ldots \psi'_{K+L+S})$, then the likelihood ratio can be formed and results in the following

$$\mathcal{L}(\theta) - \mathcal{L}(\psi) = \sum_{t=S+1}^{T} (\theta - \psi)' \mathbf{Z}_t - \sum_{t=S+1}^{T} \log \frac{c(\mathbf{Y}_{t-1} \ldots \mathbf{Y}_{t-S}; \theta)}{c(\mathbf{Y}_{t-1} \ldots \mathbf{Y}_{t-S}; \psi)}$$

[1] This establishes that a Markov Random Field can be characterized equivalently as a Gibbs distribution.

The ratio of the two, unknown normalizing constants need not be evaluated at every iteration but instead has the expectation $E_\psi\left[\frac{\exp(\theta'\mathbf{Z}_t)}{\exp(\psi'\mathbf{Z}_t)}\right]$.

A Monte Carlo estimator for this is given as

$$M^{-1}\sum_{m=1}^{M}\exp\left((\theta-\psi)'\mathbf{Z}_t^m\right)$$

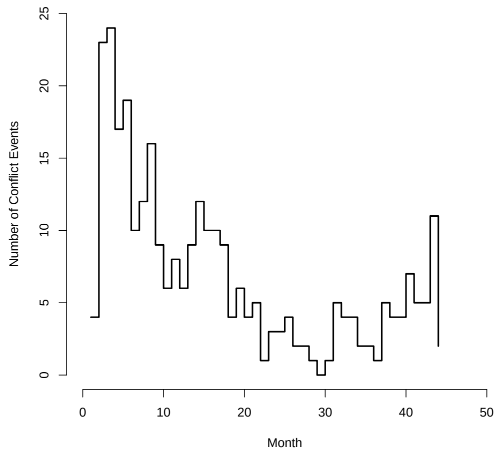
with $\mathbf{Z}_t^m$ evaluated at each of $m = 1\ldots M$ Monte Carlo samples for $\mathbf{Y}_t$. A Gibbs sampler defined by the full conditional distribution generates the Monte Carlo samples. At this point the likelihood can be maximized to obtain the MLE estimates of parameters. Typically, pseudo-likelihood estimates are used to provide an estimate of ψ, but other choices may be preferable in different scenarios.

In short, the estimation strategy is one that builds on the Geyer and Thompson (1992) insight, and creates Monte Carlo samples using a Gibbs sampler. These are then used to calculate the maximum likelihood estimates. This approach has been widely employed, and serves as the basis for Ward and Gleditsch (2002), for example. What is unique about this approach is that a temporal dimension has been incorporated, simultaneously. The temporal dimension is necessary to look more closely at the spread of events, their plausible contagion effects, and other aspects of diffusion. However, this has been dealt with typically in the social sciences by temporally lagging the spatial effects as a way of decontaminating the contemporaneous effects (Beck, Gleditsch and Beardsley, 2006). Beck et alia (2006) explore this strategy for continuous variables. Allowing one to correctly specify both temporal and contemporaneous spatial effects in discrete models improves our ability to examine contagion and diffusion processes for events that may trigger other events, which remains as one of the major hypotheses in the conflict literature (Buhaug and Gleditsch, 2008).

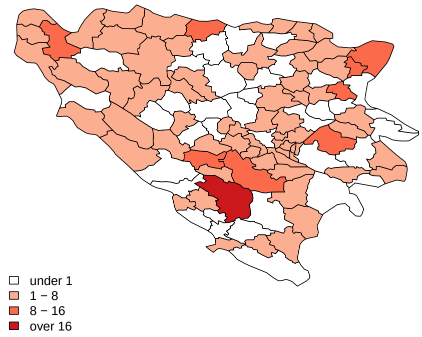
2. Application: Conflict Diffusion in Bosnia

The war in Bosnia serves as an application of the spatial-temporal model to the spread of conflict. We chose Bosnia because detailed conflict data is available at both a high spatial and temporal resolution. The 109 pre-war municipalities in Bosnia according the 1991 census (Petrovic, 1992) constitute the spatial units of analysis for our study. The boundaries of these units were digitized as a GIS shapefile, which allowed for the spatial aggregation of conflict events (see below). Our study uses months as the temporal unit of analysis. We start with March 1992 as the first month of the war, and analyze events through October 1995. This results in a sample of $N = 4,796$ (109 municipalities $\times$ 44 months). In the following, we describe the variables included in the model, and how they were computed.

2.1. **Conflict.** Data on conflict is taken from ACLED, the Armed Conflict Location and Events Dataset (Raleigh and Hegre, 2005). ACLED lists all reported confrontations between war participants along with their spatial coordinates. We aggregate ACLED events by municipality and month: specifically, we code the conflict variable as "1" if at least one ACLED event occurred in the respective unit and month, and "0" otherwise. Of the $4,796$ cases, 301 are conflict cases (6.3%). Figure 1(a) shows the conflict activity over time (aggregated by month over all municipalities). Violence erupted in two waves, one starting at the beginning of the war, and the



(a) Total number of events over time



(b) Spatial distribution of events

Figure 1. Number of conflict events per month, aggregated over all municipalities (a) and over the entire duration of the war (b).

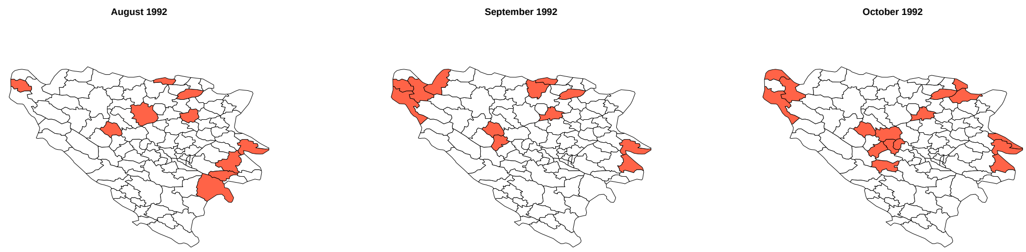


Figure 2. Diffusion of conflict, plotted over a three-month period.

second one occurring in the last 10 months before the end of the war. Figure 1(b) displays the spatial extent of violence, aggregated over the entire conflict. Conflict activity appears to be clustered in space, with a lot of violence occurring around the cities of Mostar (dark red, in the South) and Sarajevo. Also, border regions seem to be particularly affected by conflict.

The above figures are only partially able to reveal the spatial-temporal patterns of conflict in the data. Figure 2 shows the occurrence of conflict in three subsequent months (August–October 1992). There is clear evidence of conflict clustering: Take for example the municipalities in the northwest of the country. Starting with a single affected unit in August 1992, conflict spreads to the adjacent municipalities in the following months. A full spatial-temporal animation of the conflict data can be obtained from our web site.[2] Alternatively, the WarViews tool (Weidmann and Kuse, 2009) can be used to time-animate conflict data from ACLED.

[2]See `http://cederman.ethz.ch/~nilsw/bosnia/animation1.pdf`. Note that only Adobe Reader or Adobe Acrobat can play the animation. Start the sequence by clicking on the map.

2.2. **Population.** More populous municipalities should see more conflict. We therefore include the logged number of people in a unit according to the 1991 census (Petrovic, 1992) as a control variable. Municipality population varies from 4172 people (Lubinje) to 196186 (Banja Luka), with a mean of 40161. Note that Sarajevo is divided into multiple districts and therefore does not constitute the largest municipality. We expect a positive effect of population on the likelihood of conflict.

2.3. **Ethnic composition.** Ethnic composition has been shown to be a significant predictor of the location of violence (Weidmann, 2008). The war in Bosnia evolved along hardened ethnic lines, and the local ethnic composition of a municipality played a major role in making it more susceptible for violence. The 1991 Yugoslav census assigned people to different ethnic groups (Bosniaks, Serbs and Croats) and therefore can be used to calculate an index of ethnic fractionalization according to the frequently used Herfindahl concentration formula (Taylor and Hudson, 1972). In our sample, the index ranges from 0.001 to 0.66, with a mean of 0.42. We assume that higher values of ethnic fractionalization should be related to higher conflict susceptibility and therefore expect a positive effect.

2.4. **Border locations.** In the visual inspection of the conflict data presented above, it seems that border locations are particularly affected by violence. In fact, the Serb and Croat military forces moved in across the eastern and northern borders of Bosnia, which should put the border municipalities at a higher risk of conflict. We therefore include a dummy variable that takes the value of "1" for each of the 34 border municipalities.

2.5. **Elevation.** The literature on civil war has argued that mountainous regions should be more prone for conflict, since rebels can hide from government forces (Fearon and Laitin, 2003). A similar logic might apply in the Bosnian case. It will be more difficult to gain control over mountainous regions, since the opponent can more easily escape defeat. We include a measure of the average elevation of a municipality and expect a positive effect on conflict risk. Territorial elevation values are taken from the GTOPO30 dataset (US Geological Survey, 2007) are were aggregated to the municipality level using GIS software. The resulting variable ranges from 82.62 to 1283.69, with a mean of 647.92.

3. Estimation Results

We compare the MCMC estimation results obtained with our `streg` package (see Appendix A) to the ones obtained using maximum pseudo-likelihood (MPL) estimation. For the MCMC estimation, we used 11,000 Gibbs sampling iterations with a burnin period of 1000. In both models, we use temporal lags up to the second order. Table 1 reports the results.

We see that the effects of the variables on conflict risk are largely as expected: More populous units face a higher risk of conflict, and so do ethnically diverse units. Proximity to the country border increases the likelihood of violence. Conflict is also more likely in mountainous terrain. The results show the strong spatial and temporal dependence in the data. The spatial lag gets a positive and strongly significant coefficient. In other words, the occurrence of violence in a municipality spurs violence nearby – this confirms our visual impression from the conflict data presented above.

	MPL estimation		MCMC estimation	
	$\hat{\psi}$	$\sigma_{\hat{\psi}}$	$\hat{\theta}$	$\sigma_{\hat{\theta}}$
(Intercept)	-12.0819	1.2953	-11.8398	1.2504
Population	1.5725	0.2674	1.5051	0.2556
ELF	1.1800	0.5572	1.1118	0.5224
Border	0.5704	0.1707	0.6657	0.1646
Mountains	0.0008	0.0003	0.0009	0.0002
Conflict (Spatial lag)	0.8074	0.0761	0.8341	0.0779
Conflict (t-1)	1.9605	0.1669	2.0052	0.1580
Conflict (t-2)	0.7210	0.1884	0.8064	0.1772

Table 1. Results for Pseudo-likelihood and MCMC estimation.

Also, violence displays a strong temporal dependence, which becomes visible in the positive and significant time lags.

Comparing the results from the MPL and MCMC estimations, we note that the differences are not large. As reported in Ward and Gleditsch (2002), we find that the coefficients estimated by MPL are close to those obtained with the MCMC model. Also in line with Ward and Gleditsch's findings, we see that the standard errors estimated by MCMC tend to be smaller, even though in our application this does not alter the substantive conclusions. In the past, due to the comparably high computational complexity of the MCMC procedure, researchers have often resorted to fast MPL estimation. However, the introduction of our `streg` package makes it possible to get MLE estimates quickly.[3] In essence, whether MCMC estimates differ from MPL estimates depends on the respective application, and we cannot tell unless we use both estimation methods. Our package enables researchers to use the appropriate estimation for spatially dependent outcomes and obtain results fast.

4. Evaluating the Predictions of the Model

This section describes our efforts to evaluate the predictive accuracy of the spatial-temporal model. We use the model specification with two time lags presented above. As a result of omitting the first two months in the dataset, we end up with 42 months. Predicting a rare event as in our case (273 positive cases out of $4,578$) is a difficult undertaking, especially since our covariates do not display variation over time. Instead, the spatial and temporal lags in our model will carry the major burden of predicting the temporal variation of conflict. The predictions we report here are in-sample predictions, since the model was built on the full set of data.

4.1. **Predictions Based on the Observed Conflict Outcomes.**

We start by examining the predicted values of the model based on the observed values for both the spatial and temporal lags. More precisely, we compute the predicted probability $\hat{p}_{i,t}$ of unit i at time t as

$$\hat{p}_{i,t} \quad = \quad \text{logit}^{-1}\left(\sum_{k=0}^{K}\hat{\theta}_k X_{k,i,t} + \hat{\theta}_{K+1}\sum_{j\in N_i} Y_{j,t} + \sum_{s=1}^{2}\hat{\theta}_{K+1+s} Y_{i,t-s}\right)$$

[3] The above presented model took about a minute to run on an Intel Core Duo 1.6 GHz CPU.

with observed values $Y_{j,t}$. The predicted probabilities we obtain are generally fairly small. Figure 3 shows a plot of their density. Based on the inspection of the figure, we use a cutoff value of 0.2 for positive classifications. The complete truth table is shown below (Table 2).

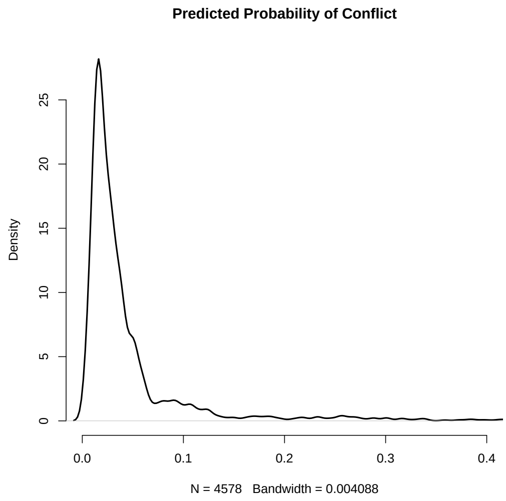


FIGURE 3. Density of the predicted conflict probabilities.

	Observed 0	Observed 1
Predicted 0	4118	155
Predicted 1	186	119

TABLE 2. Truth table for the predictions generated with the spatial-temporal model. Threshold for positive classifications: 0.2.

Note that the procedure presented above is not an accurate assessment of the model predictions due to the simultaneity of the observations across space. Below, we present results obtained using the approach described in Ward and Gleditsch (2002). Essentially, the idea is to use the estimated $\hat{\theta}$ of the MCMC model to generate predicted probabilities of conflict by means of simulation. Rather than using the *observed* conflict outcomes as spatial lags during the simulation, we update the predicted value of a municipality based on the *simulated* values of its neighbors. To initialize the simulation, we use the observed conflict outcomes as starting values and update each of the municipalities 100 times in a random sequence. During the simulation, we assign a simulated conflict value of "1" to a unit if the computed conflict probability exceeds 0.2. The final conflict predictions are the values obtained after the number of simulation iterations.

Whereas the problem of simultaneity can be resolved using the simulation approach described in Ward and Gleditsch (2002), for the spatial-temporal model, we also have to deal with the time dimension. Basically, when computing the predicted probability $\hat{p}_{i,t}$ of unit i at time t, we can use either the *observed* temporal lags, or the *simulated* ones obtained for the previous time steps. We describe each of the two approaches below.

4.2. **Predictions Using the Observed Conflict History.** Our first simulation approach uses the observed temporal lags of the dependent variable, but the simulated spatial lags. More formally, we compute the predicted probability $\hat{p}_{i,t}$ as

$$\hat{p}_{i,t} = \text{logit}^{-1}\left(\sum_{k=0}^{K} \hat{\theta}_k X_{k,i,t} + \hat{\theta}_{K+1} \sum_{j \in \mathcal{N}_i} \tilde{Y}_{j,t} + \sum_{s=1}^{2} \hat{\theta}_{K+1+s} Y_{i,t-s}\right)$$

where $\tilde{Y}_{i,t}$ denotes the simulated conflict value obtained for the same time step. Compared to the classifications we received when using observed values for both the spatial and temporal lags (see Table 2 above), the predictive accuracy of the model improves considerably. However, the precision of the predictions suffers. Out of 441 cases classified as positive, slightly less than half (141) are real conflict cases (see Table 3).

	Observed 0	Observed 1
Predicted 0	4004	133
Predicted 1	300	141

Table 3. Truth table for the predictions generated using simulation with observed temporal lags.

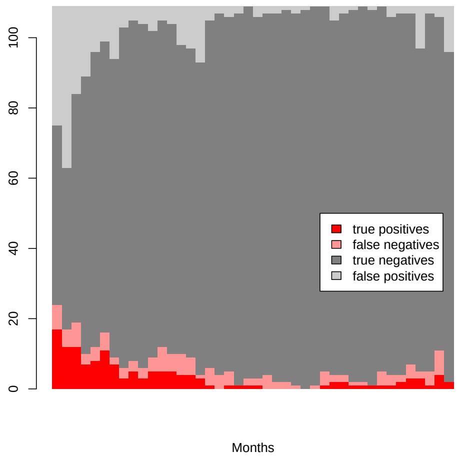
Whereas the above truth table displays aggregate statistics over the entire study period, it is interesting to trace the predictive accuracy as it develops over time. Figure 4(a) plots the numbers of true and predicted conflict cases. Red and grey colors indicate true conflict and non-conflict cases, respectively. Dark colors show the number of correct classifications. We see that there is almost no variation of the accuracy of classification over time: the number of correctly predicted conflict cases remains low throughout the entire study period.

4.3. **Predictions Using the Simulated Conflict History.** We conduct are more difficult prediction task using simulated values for both the spatial and temporal lag. Essentially, the predicted values are computed according to

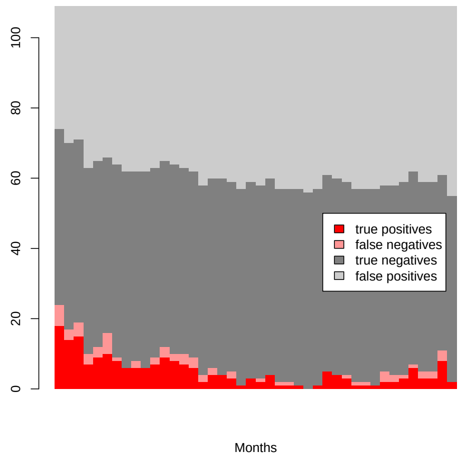
$$\hat{p}_{i,t} = \text{logit}^{-1}\left(\sum_{k=0}^{K} \hat{\theta}_k X_{k,i,t} + \hat{\theta}_{K+1} \sum_{j \in \mathcal{N}_i} \tilde{Y}_{j,t} + \sum_{s=1}^{2} \hat{\theta}_{K+1+s} \tilde{Y}_{i,t-s}\right)$$

where $\tilde{Y}_{i,t}$ denotes again the simulated conflict value of unit i at time t. We expect this approach to yield worse results than the previous one, since wrong classifications will propagate through the system over time, whereas in the previous approach, they could not due to the use of the real conflict outcomes for the temporal lags. As expected, the results are worse compared to the previous simulation. We are able to identify 207 conflict cases correctly, but at the same time, we wrongly classify 2025 cases as positive.

Figure 4(b) shows again a visual illustration of the results. Because of the temporal propagation of error, we would expect the predictive performance to worsen as the simulation evolves over time. Indeed, there seems to be decrease in the correctly predicted positive instances over time, albeit a weak one.



(a) Classifications obtained using observed temporal lags



(b) Classifications obtained using simulated temporal lags

Figure 4. Classification results for the simulated predictions, for each of the 42 months (along the x-axis). Dark and light red: true conflict cases. Dark and light grey: true non-conflict cases. The dark colors indicate the correctly classified instances.

	Observed 0	Observed 1
Predicted 0	2279	67
Predicted 1	2025	207

Table 4. Truth table for the predictions generated using simulation with simulated temporal lags.

5. Conclusion

In this paper, we have presented an attempt add a time dimension spatial regression models of conflict. To this end, we developed a spatial-temporal regression model following recent work (Zhu, Zheng, Carroll and Aukema, 2008) and applied it to the diffusion of conflict during the Bosnia War, using monthly conflict data at the level of municipalities. The results show that conflict displays both a strong spatial and temporal dependence. These results were obtained using our `streg` package for $\mathcal{R}$ that facilitates the estimation of spatial-temporal regression models. This package is currently at the development stage, but will later be extended to allow for the estimation of different spatial-temporal regression models and their predictive evaluation.

We provided an extensive assessment of the model's predictive performance, with mixed results. Using simulation to incorporate the simultaneity of conflict observations over space, we derived conflict predictions and compared them to the observed outcomes. The challenge during these simulations is the inclusion of the time component. Ideally, we want to initialize the simulation with a given set of cases and then project its predictions both across space and time. We began this effort in our second simulation exercise, but different projection methods are under development. Obviously, this approach is difficult, since prediction errors made early in the process will propagate through the entire history of the simulation. We presented another simulation approach that inhibits the temporal propagation of error by relying on the observed conflict history of a unit when computing the conflict prediction. In general, however, to further evaluate the model, we need to do out-of-sample prediction within the context of a simulation that is based on the MCMC samples generated from the full conditionals.

Appendix A. R-package `streg` for Space-Time Autologistic Models

The `streg` package provides an implementation of Zhu et al.'s spatial-temporal regression model (Zhu, Zheng, Carroll and Aukema, 2008). Whereas it is possible to implement the model estimation directly in the $\mathcal{R}$ statistical package, in doing so one would incur significant performance costs. The reason is that $\mathcal{R}$ is optimized for computation on vectors of data. Repeated retrievals of individual records from the data matrix – just as we do when running the Gibbs sampler – is not very efficient and leads to very long execution times of the algorithm.

For that reason, we chose to implement the model estimation in the Java programming language. In contrast to other programs written e.g. in C or C++, Java programs can be executed under almost any operating system, which prevents the developer from having to distribute a particular piece of software for different operating systems. Using the `rJava` package[4] for $\mathcal{R}$, we interface from $\mathcal{R}$ to Java. The basic input checking is done in $\mathcal{R}$, and if successful, the data are sent to a newly created Java object. Once the computation is done, the $\mathcal{R}$ system retrieves the results and formats the output. Currently, the package only offers the `stlogit` command described below. However, we are planning to expand the package's functionality to incorporate spatial-temporal models for alternative dependent variables. The syntax of the `stlogit` command is as follows:

```
> stlogit(model, data, unitvar, timevar, numtimelags, weights,
iterations, burnin)
```

- `model`: the specification of the regression model in standard $\mathcal{R}$ syntax, with spatial and temporal lags omitted, as they will be automatically computed by the package.
- `data`: the data frame
- `unitvar`, `timevar`: Unique identifiers for the spatial units and the time periods.
- `numtimelags`: The desired number of time lags. Currently, the model always computes spatial lags of order 1.
- `weights`: The weights matrix.
- `iterations`: The total number of Gibbs iterations (including the burnin period).
- `burnin`: The number of burn-in periods.

[4]see http://www.rforge.net/rJava/

References

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

Besag, Julian E. 1972. "Nearest-neighbour Systems and the Auto-logistic Model for Binary Data." *Journal of the Royal Statistical Society, Series B, Methodological* 34:75.

Besag, Julian E. 1974. "Spatial Interaction and the Statistical Analysis of Lattice Systems (with Discussion)." *Journal of the Royal Statistical Society, Series B, Methodological* 36:192–225.

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

Diggle, Peter J., Rana A. Moyeed and Jonathan A. Tawn. 1998. "Model Based Geostatistics, with discussion." *Applied Statistics* 47:299–350.

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

Geyer, Charles J. and Elizabeth A. Thompson. 1992. "Constrained Monte Carlo Maximum Likelihood for Dependent Data, (with Discussion)." *Journal of the Royal Statistical Society, Series B, Methodological* 54:657–699.

Gleditsch, Kristian S. 2000. "War and Peace in Space and Time: The Role of Democratization." *International Studies Quarterly* 44(1):1–29.

Gleditsch, Kristian S. and Michael D. Ward. 2001. "Measuring Space: A Minimum Distance Database and Applications to International Studies." *Journal of Peace Research* 38(6):749–768.

Gumpertz, Marcia L., Chi-tsung Wu and John M. Pye. 2000. "Logistic Regression for Southern Pine Beetle Outbreaks with Spatial and Temporal Autocorrelation." *Forest Science* 46(1):95–107.

Hoeting, Jennifer A., Molly Leecaster and David Bowden. 1999. "An Improved Model for Spatially Correlated Binary Responses." *Journal of Agricultural, Biological and Environmental Statistics* 5(1):102.

Huffer, Fred W. and Hulin Wu. 1998. "Markov Chain Monte Carlo for Autologistic Regression Models with Application to the Distribution of Plant Species." *Biometrics* 54(2):509.

Petrovic, Ruza. 1992. "The National Composition of Yugoslavia's Population, 1991." *Yugoslav Survey* 33(1):3–24.

Raleigh, Clionadh and Haavard Hegre. 2005. "Introducing ACLED: An Armed Conflict Location and Event Dataset." Paper prepared for the Disaggregating the Study of Civil War and Transnational Violence Conference, San Diego, CA.

Taylor, Charles and Michael C. Hudson. 1972. *World Handbook of Political and Social Indicators*. New Haven: Yale University Press.

US Geological Survey. 2007. "GTOPO30 Digital Elevation Model.". Available at http://edc.usgs.gov/products/elevation/gtopo30/gtopo30.html.

Ward, Michael D. and Kristian Skrede Gleditsch. 2002. "Location, Location, Location: An MCMC Approach to Modeling the Spatial Context of War and Peace." *Political Analysis* 10(3):244–260.

Weidmann, Nils B. 2008. "Violence and the Changing Ethnic Map: The Interdependence of Territory and Conflict in Bosnia." Paper prepared for the Annual

Convention of the International Studies Association, San Francisco.

Weidmann, Nils B. and Doreen Kuse. 2009. "WarViews: Visualizing and Animating Geographic Data on Civil War." *International Studies Perspectives* . Forthcoming.

Zheng, Yanbing and Jun Zhu. 2008. "Markov Chain Monte Carlo for a Spatial-Temporal Autologistic Regression Model." *Journal of Computational & Graphical Statistics* 17(1):123–137.

Zhu, Jun, Jakob G. Rasmussen, Jesper Møller, H. Aukema, Brian and Kenneth F. Raffa. 2008. "Spatial-Temporal Modeling of Forest Gaps Generated by Colonization From Below- and Above-Ground Bark Beetle Species." *Journal of the American Statistical Association* 103(481):162.

Zhu, Jun, Yanbing Zheng, L. Carroll, Allan and H. Aukema, Brian. 2008. "Autologistic Regression Analysis of Spatial-Temporal Binary Data via Monte Carlo Maximum Likelihood." *Journal of Agricultural, Biological & Environmental Statistics* 13(1):84–98.

Nils B. Weidmann: ETH Zurich, International Conflict Research, Seilergraben 49, 8092 Zurich, Switzerland
E-mail address: `weidmann@icr.gess.ethz.ch`

Michael D. Ward: Department of Political Science & Center for Statistics and Social Science, University of Washington, Seattle, Washington, USA, 98195-3530
E-mail address: `mdw@u.washington.edu`