

A roadmap to realistic computational models of civil wars

Lars-Erik Cederman and Luc Girardin

International Conflict Research
ETH Zurich
8092 Zurich, Switzerland
{cederman, girardin}@icr.gess.ethz.ch

Introduction

Civil wars arguably pose the most serious problem to security in the contemporary world. Since the end of World War II, internal conflicts have become much more frequent than wars among states (Sollenberg and Wallensteen 2001). Accordingly, it does not come as a surprise that most conflict analysts have increasingly come to shift their research toward civil wars (Lake 2003). There is a vibrant, interdisciplinary literature that applies sophisticated research methods, including statistics and formal modeling, to such conflicts (Sambanis 2004).

Similarly, agent-based modelers have traditionally analyzed interstate warfare, but they have been slower to embrace civil wars as their research focus. This paper reports on our own efforts to build more accurate computational models of civil wars, and sketches a computational research program for this purpose. We propose solutions to the conceptual and methodological obstacles that stand in the way of progress. In particular, we suggest ways to integrate data generated in geographic information systems (GIS) with agent-based models.

Traditional models of interstate warfare

Geopolitics belongs to the core research agenda of agent-based modelers. Inspired by Realism, the dominant paradigm of International Relations, such modeling efforts typically feature states as autonomous territorial actors embedded in a decentralized system that exhibits balance-of-power equilibria. In a pioneering paper, Bremer and Mihalka (1977) introduced such a model of geopolitical competition. It features state-like organizations with dynamic borders that grow through conquest.

Building on Bremer and Mihalka's original model, Cusack and Stoll (1990) presented a book-length study that analyzes the internal consistency of the realist paradigm. This study includes a systematic set of computational experiments featuring an extended set of decision-making rules. Civil wars were also modeled, but because of the complexity of the model and the limited computational resources available to the authors, the book does not contain any systematic replications of internal conflict.

The GeoSim framework

This research tradition became an important source of inspiration for the GeoSim project. Introduced by Cederman (1997) and modeled from scratch, GeoSim is a family of agent-based models that is based on a dynamic network of interstate relations superimposed on a square grid. In the basic model, all interactions are local, between adjacent states. Each state capital can absorb and dominate a number of provinces in a perfectly Hobbesian fashion. Moreover, their borders are sharply defined. Finally, they derive their power from the number of provinces they control -- thus, the larger a state is, the more powerful it is.

The main thrust of the GeoSim research program is to study interstate warfare as a consequence of geopolitical changes affecting the boundaries of states (Cederman 2002). Within this context, it is possible to apply the framework to study the effect of defensive alliances and technology (Cederman 1997, Chap. 4), democracy (Cederman 2001), and democratization (Cederman and

Gleditsch 2004) on interstate conflict processes, as well as the nature of such processes, including war-size distributions based on casualty levels (Cederman 2003).

Another stream of research based on the GeoSim framework opens the black box of the state by allowing provinces to stage rebellions against the central power of the state. Focusing on nationalist challenges to the capital within a single state, Cederman (1997: Chap. 8) presents an early attempt to relax the assumption of unitary states. This model introduces a straight-forward center-periphery logic to the model that follows the hierarchical structure of the states. Here the provinces manage to overcome the power balance in favor of the center by subscribing to nationalist platforms. It is shown that the more acute the oppression exerted by the center, the more likely it is that the peripheral actors manage to find a least common denominator based on a “thin” rather than a specific identity.

The nationalist insurgency model

Going a step beyond this simple research design, Cederman (2006) proposes another variation on the GeoSim structure, the nationalist insurgency model (NIM), in which each simulation run consists of an entire state system, albeit with fixed state borders. This model was developed to reconstruct the mechanisms that drive conflict in center-periphery relationships within states. In an oft-cited study, Fearon and Laitin (2003) suggest that weak states characterized by rough terrain are especially conflict prone because they are incapable of controlling their territories, thus opening a window of opportunity to rebels. However, their findings do not lend any support to hypotheses that connect ethnic diversity with civil wars.

Challenging this claim, Cederman (2006) manages to reproduce Fearon and Laitin’s results on state size and terrain while insisting that ethnicity and nationalism play a key role in generating conflict. This means that the NIM features a center-periphery logic with a cultural dimension rather than being merely materialist. The computational experiments show that civil wars may result where the state is incapable of full cultural penetration resulting in incomplete nationalist mobilization and peripheral collective action based on common identities that challenge the control of the center.

In order to generate findings on nationalist insurgencies, the NIM features a number of additions to the standard interstate framework. Instead of including merely one “layer” pertaining to relations among states, the current model introduces an explicit model of terrain as well as a multi-dimensional cultural map similar to the one proposed in Axelrod (1997: Chap. 7). Inspired by Holland’s (1995) schema representation, national identities are modeled as computational coalitions defined in terms of a culture string with “wild cards” that is superimposed on top of the cultural landscape.

Figure 1 shows a 3D snapshot of the model that reveals the rugged physical landscape. Here state borders appear as lines and the state capitals as half spheres. Both capitals and provinces are marked by colored disks if they possess a national identity. Rebellions are shown as vertical “needles”. Interestingly, the projections suggest that most of the fighting takes place in mountainous areas.

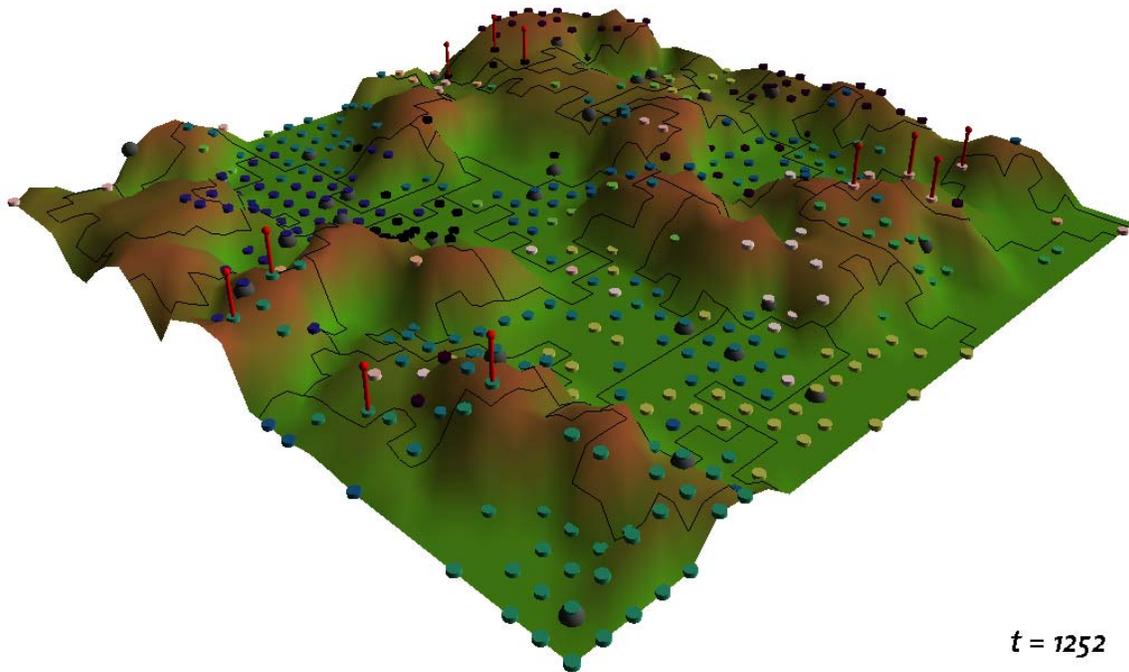


Figure 1. A 3D-projection of the nationalist insurgency model showing rebelling provinces as red needles.

This representation goes beyond models of culture as fixed and immutable properties that influence behavior (Epstein 2002) or as endogenous vectors whose traits all matter (Axelrod 1997, Chap. 7). Identities, however, are more selective, because only politically relevant aspects of culture enter into the power calculus. Allowing for such variation, Lustick's (2000; 2002) agent-based models ABIR and PS-I feature endogenous repertoires composed of sets of identities, but the component identities stand in no specific relationship to each other and thus do not describe a coherent cultural space. Moreover, these and almost all other models of this type fail to provide an explicit representation for formal political organizations or terrain.

While the NIM offers considerable flexibility to represent complex phenomena, it is entirely heuristic. Other than reproducing similar macro results as those found by Fearon and Laitin (2003), no attempt has been made to calibrate the key objects to real-world conditions. This raises the question of whether it is possible to build agent-based models that conform more closely to real-world phenomena.

Toward more accurate agent-based models of civil wars

In the following, we describe the steps that we have taken more recently to put our modeling activities on a more stable empirical footing. As was the case in the NIM, our research focuses entirely on ethno-nationalist civil wars following a center-periphery logic, thus setting aside other types of wars that are not ethnic or that involve no hierarchy, such as communal riots.

Beyond fractionalization

In particular, our previous computational research suggests that it is important to determine the relative power of the ethnic groups and their geographic location. Therefore we launched two data projects that serve to investigate these issues. The first one, reported in Cederman and Girardin (2005) replaces the conventional index of ethnic fractionalization with a measure that we call N^* . Directly drawing on the center-periphery of the NIM, this index measures the extent to which peripheral ethnic groups are excluded from state power in states by comparing their demographic share to that held by the “ethnic group in power” (EGIP). For the countries that have so far been coded, we have obtained strongly significant results when we regress ethnic civil wars on N^* while controlling for the variables used by Fearon and Laitin (2003).

Geo-referencing of ethnic groups

This result confirms that the center-periphery logic of the NIM is sound. However, it says little about the geographic dimension. For this reason, we initiated a second project that puts real empirically observed ethnic groups on the map. This data project that serves to geo-reference ethnic groups around the world (Cederman, Rød and Weidmann 2006). Relying on maps and data drawn from the classical Soviet *Atlas Narodov Mira*, we use geographic information systems (GIS) to represent the groups as polygons. The Atlas has several strengths: it is complete and carefully researched, it relies on a uniform group list that is valid across state borders, and it provides high-quality maps. Among the weaknesses, it should be stressed that the Atlas is based on the situation in the 1960s and thus clearly outdated compared to the current situation. However, in most cases, ethnic settlement patterns exhibit considerable inertia, so it seems reasonable to use this dataset as a starting point. It is possible to reconstruct population shares with GIS procedures. Figure 2 provides a snapshot of geo-coded ethnic groups in the former Yugoslavia.



Figure 2. Geo-coded map of the former Yugoslavia with ethnic groups represented by polygons.

Modeling ethnic conflict in center-periphery dyads

In a related project together with Halvard Buhaug and Jan Ketil Rød, we have already started to analyze the geo-coded data on ethnic groups with logistical regression. As in the computational framework, we focus on relation-specific causes of ethnic conflicts. The goal is to disaggregate both ethnicity and conflict to the level of explicitly geo-coded center-periphery dyads. This is an important goal, because, so far, conflict data has typically been recorded at the country level (Sambanis 2003). Whereas this is a satisfactory simplification in small countries with few conflict groups, the results can be seriously misleading in the case of large countries, such as Russia. The political, cultural and geographic conditions pertaining to Chechnya do not necessarily apply elsewhere in the country.

Focusing specifically on ethno-nationalist conflicts pitting peripheral ethnic groups against central governments, this analysis allows us to measure the center-periphery power balance as terrain-adjusted demographic proxies. In addition to establishing a link between the dyads and conflict data, it is essential to pinpoint the location of the ethnic groups themselves.

Table 1 presents our preliminary regression analysis. In this case, the observations are 792 ethnic center-periphery dyads where the center is represented by an EGIP. The sample is limited to Eurasia and North Africa as in Cederman and Girardin (2005). Our dependent variable is coded as one if the dyad experienced an ethnic war during the period from 1945 through 1999 and zero otherwise. The table reveals that the power balance, measured as the peripheral group's share of the dyadic resources, has a strong, highly significant effect. Furthermore, our measure of logistical obstacles, which is operationalized as the product of the ethnic groups' average distance from the capital multiplied by its share of mountainous terrain, also reaches significance, although merely at the 0.05 level. Finally, we control for logged GDP per capita, measured at the country level, which is not significant. The latter result stands in contrast to otherwise robust findings at the country level, indicating that GDP has a strong negative impact on conflict.

Table 1. Preliminary logit analysis of onset of ethnic civil wars

Logit estimates				Number of obs	=	792
Log pseudo-likelihood = -216.80482				Pseudo R2	=	0.0746
				(standard errors adjusted for clustering on cow)		

			Robust			
conflict		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]

power balance		7.375405	1.850482	3.99	0.000	3.748526 11.00228
logistic		.1354501	.0648003	2.09	0.037	.0084438 .2624564
logGDP		-.0652103	.2006814	-0.32	0.745	-.4585386 .3281181
_cons		-2.606658	1.780366	-1.46	0.143	-6.096112 .8827954

By disaggregating the conventional country-level analysis to the level of center-periphery dyads, we are now in a position to formulate and validate hypotheses that feature explicit, group-level mechanisms and motivations. However, the empirical analysis has so far been entirely static. In order model conflict processes and their constitutive mechanisms dynamically, it is necessary to return to computational modeling.

GROWLab

As we have seen, the NIM creates an artificial world that can be used for exploratory modeling. The next step toward higher realism requires a tighter coupling of the already collected GIS data and our computational tools (Brown et al. 2005). Rather than continuing to build on GeoSim, we therefore decided to create the next generation of geopolitical simulation framework in collaboration with Nils Weidmann. The result is GROWLab (Geographic Research on War Laboratory), which is a software toolbox to facilitate the modeling, simulation, analysis, and validation of complex social processes, with a special focus on state and civil violence. GROWLab is designed as a collection of independent software components, which are expected, when finalized and put together, to bring the development of agent-based simulations to the next level of complexity and realism.

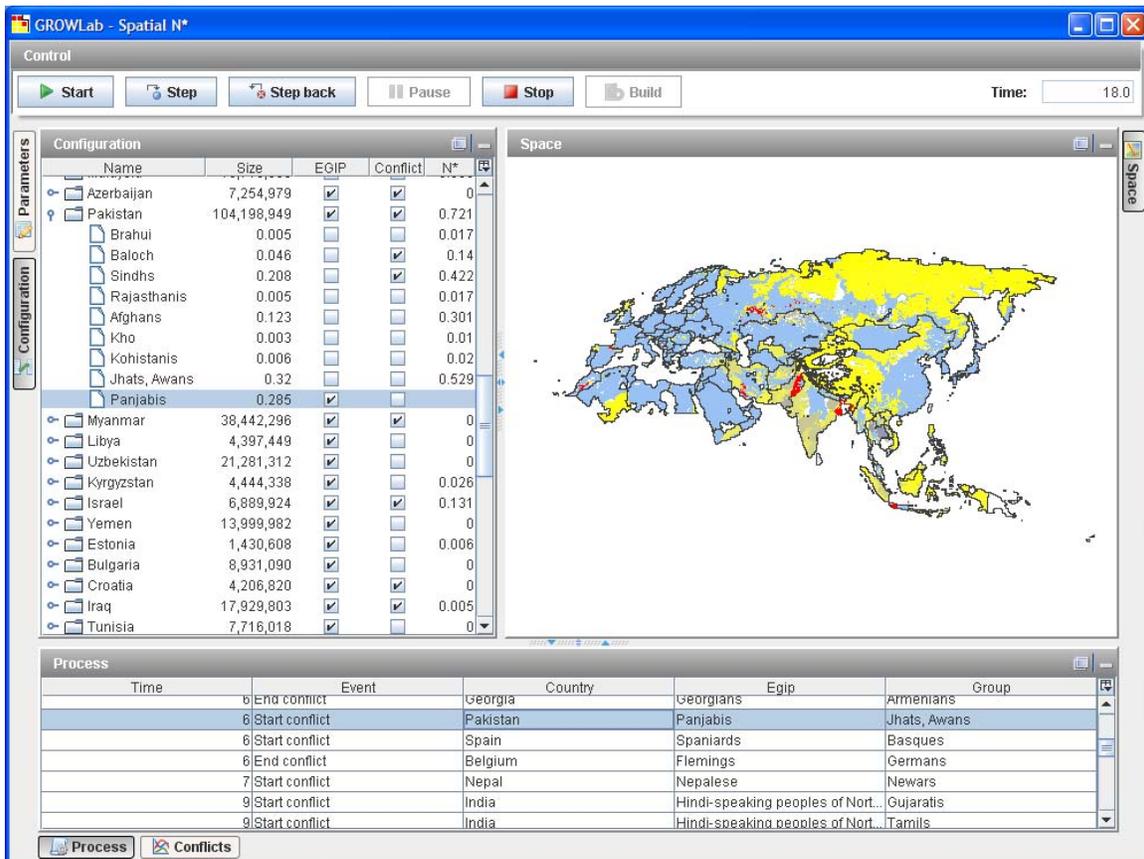


Figure 3. Our model aiming at simulating both the onset and duration of civil wars.

GROWLab is used to calibrate the model with empirical facts and geographical features, including the location of ethnic groups and their political relationships. Multiple coordinated views allow understanding different aspects of the simulation, from the political configuration of the system, its representation in space, to the unfolding of various processes in place.

The specific aim of this work is to develop a Java class library as a series of toolboxes. Initially, four core components are planned to be developed to support:

- seeding of the model with empirical facts, including GIS-based data, to calibrate the environments and mechanisms to the appropriate level of realism;

- effective modeling of complex network and hierarchical relationships between model actors and the efficient scheduling of their interactions;
- execution of large number of simulation runs on a grid made of many independent computers to test the sensitivity of the models;
- statistical and visual analysis of the state of the system, as well as the unfolding of the processes over time.

Figure 3 shows the current state of our attempt to capture the link between ethnicity and civil violence. We use it to test specific causal mechanisms that connect different types of ethno-nationalist configurations with the outbreak and extension of internal conflict. Our goal is to explain such phenomena as integrated parts of spatiotemporal macro processes.

Future research agenda

Equipped with our new geographic data and GROWLab, we intend to craft a new series of models that go beyond existing statistical modeling of civil wars. Computational techniques are especially helpful when it comes to capturing spatiotemporal processes, identity politics, and systemic effects linking several states into a system. In these cases, non-computational methods usually fail to offer sufficient flexibility of representation (Cederman 1997).

For example, extensive quantitative research exists on the outbreak and duration of civil wars respectively, but there are no integrated models that succeed in bringing these conflict dimensions together. Moreover, agent-based models are likely to provide valuable insights into how to explain war-size distributions. Following the regularity of interstate wars (Cederman 2003), recent findings suggest that the casualty numbers of certain civil wars and terrorist events are power-law distributed (e.g. Johnson et al. 2005). It remains to be seen if these findings are robust, but whether the empirical distribution is a power law or some other skew distribution, it can be expected that macro-level results of this type will help us in the search for causal mechanisms that drive internal conflict.

Empirically supported computational modeling promises to move the literature from abstract correlations based on merely postulated micro foundations to the articulation of theoretical models featuring explicit actors and actions. Ultimately, in this way, we hope to address the important theoretical questions about whether and how ethnicity plays a role in civil wars

References

- Axelrod, Robert. 1997. *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*. Princeton: Princeton University Press.
- Bremer, Stuart A., and Michael Mihalka. 1977. Machiavelli in Machina: Or Politics Among Hexagons. In *Problems of World Modeling*, edited by K. W. Deutsch. Boston: Ballinger.
- Brown, Daniel G. et al. 2005. Spatial Process and Data Models: Toward Integration of Agent-Based Models and GIS. *Journal of Geographic Systems* 7: 1-23.
- Cederman, Lars-Erik. 1997. *Emergent Actors in World Politics: How States and Nations Develop and Dissolve*. Princeton: Princeton University Press.
- Cederman, Lars-Erik. 2001. Modeling the Democratic Peace as a Kantian Selection Process. *Journal of Conflict Resolution* 45 (August):470-502.
- Cederman, Lars-Erik. 2002. Endogenizing Geopolitical Boundaries with Agent-Based Modeling. *Proceedings of the National Academy* 99 (suppl. 3):7796-7303.
- Cederman, Lars-Erik. 2003. Modeling the Size of Wars: From Billiard Balls to Sandpiles. *American Political Science Review* 97 (February):135-150.

- Cederman, Lars-Erik. 2006 (forthcoming). Articulating the Geo-Cultural Logic of Nationalist Insurgency. In *Order, Conflict, and Violence*. Cambridge: Cambridge University Press.
- Cederman, Lars-Erik, and Luc Girardin. 2005. Beyond Fractionalization: Mapping Ethnicity onto Nationalist Insurgencies. Paper read at Disaggregating the Study of Civil War and Transnational Violence Conference, at San Diego, CA.
- Cederman, Lars-Erik, and Kristian Skrede Gleditsch. 2004. Conquest and Regime Change: An Evolutionary Model of the Spread of Democracy and Peace. *International Studies Quarterly* 48 (3):603-629.
- Cederman, Lars-Erik, Jan Ketil Rød and Nils Weidmann. 2006. Geo-Referencing of Ethnic Groups: Creating a New Dataset. Paper presented at the GROW-Net Workshop, PRIO, Oslo.
- Cusack, Thomas R., and Richard Stoll. 1990. *Exploring Realpolitik: Probing International Relations Theory with Computer Simulation*. Boulder: Lynnie Rienner.
- Epstein, Joshua M. 2002. Modeling Civil Violence: An Agent-Based Computational Approach. *Proceedings of the National Academy of Sciences* 99 suppl. 3:7243-7250.
- Fearon, James D. 2003. Ethnic and Cultural Diversity by Country. *Journal of Economic Growth* 8:195-222.
- Holland, John H. 1995. *Hidden Order: How Adaptation Builds Complexity*. Reading, Mass.: Addison-Wesley.
- Johnson, N. et al. 2005. From old wars to new wars and global terrorism. *Physics Abstracts* physics/0506213.
- Lake, David A. 2003. International Relations Theory and Internal Conflict: Insights from the Interstices. *International Studies Review* 5 (4):81-89.
- Lustick, Ian S. 2000. Agent-Based Modelling of Collective Identity: Testing Constructivist Theory. *Journal of Artificial Societies and Social Simulation* 3 (1).
- Lustick, Ian S. 2002. PS-I: A User-Friendly Agent-Based Modeling Platform for Testing Theories of Political Identity and Political Stability. *Journal of Artificial Societies and Social Simulation* 5 (3).
- Sambanis, Nicholas. 2004. Using Case Studies to Expand Economic Models of Civil War. *Perspectives on Politics* 2: 259-279.
- Sollenberg, Margareta, and Peter Wallensteen. 2001. Patterns of Major Armed Conflicts, 1990-2000. In *SIPRI Yearbook 2001. Armaments, Disarmament and International Security*. Oxford: Oxford University Press.