

Chapter 12

New Methods for Conflict Data

Will Lowe

Harvard University, Cambridge, Massachusetts, U.S.A.

This chapter sketches out some new ways to look at conflict data sets. Since political scientists have more computer power available to them than at perhaps any point in the past, the paper emphasizes methods whose principle feature is that they purchase substantive realism at the cost of more compute cycles, not more advanced statistics. Existing statistical theory is sufficient to perform much more realistic analyses than are typically performed, but it is not necessarily found in the standard location. Most of the models and methods described here can be found in other guises in the field of machine learning. Political methodology is often accused of importing techniques wholesale from other disciplines, particularly econometrics, and by introducing machine learning as another field worth mining, this paper continues a long tradition.¹

The sections below consider new approaches to two important sources of information about conflict: event data and conflict databases. Event data remain underused in international relations, in part because of the apparently narrow range of inferences they license using existing methods. Hopefully, widening the range of available methods should earn this sort of data a wider audience, and help integrate them more fully into quantitative international relations research. Conflict data sets, e.g. the State Failure, and Correlates of War datasets (Esty et al., 1998; Singer and Diehl, 1990), in contrast, are widely used to provide data for regression analysis and this application is well-understood. Trappl and colleagues have recently applied case-based reasoning methods to making predictions (see Chapter 11 in this volume). We describe how this work might itself be generalized in a way that allows it to be integrated back into regular statistical methodology using probabilistic expert systems.

¹. But it is a good one; methodology should always be about ensuring a model represents your theory, not somebody else's, so ultimately it is the theory that matters, and any model that effectively represents it is good enough, wherever it came from.

1 EVENTS DATA

Event data analysis is motivated by McClelland's observation that international interactions can be understood as the cumulative effects of large numbers of stereotypical dyadic actions, or events (McClelland, 1978; Gerner et al., 1994). Three aspects are important to this view of international relations: first, events are stereotypical because they come in a finite number of possible types. They are also relatively unambiguous. Theoretically this is because although the typology is typically hierarchical with a single root action node (e.g. IDEA²) or a set of root nodes (e.g. WEIS, McClelland, 1978; CAMEO, Gerner et. al, 2002), no leaf action can be a child of more than one parent. Thus there is no typological ambiguity as to the meaning of an event. They are also practically unambiguous because events, in contrast to the intentions of the actors involved, are in principle observable. The second aspect is that event data sets are focused around the country dyad. This contrasts with case study methods that concentrate on the multilateral relations of one actor. The third aspect is that, given their definition, the natural numerical summary of event data is the count.

Much event data analysis involves regressing sets of counts, or more usually sets of scaled counts, on each other. This practice makes event data analyses a natural target for critique, from two directions. From one side, researchers immersed in case studies and historical research object to the reduction of subtle cultural and institutional processes to a universal set of decontextualized actions and reactions. From the other side, game theorists, who may not be so averse to reductive theories, complain that the strategic choices, reputation effects, and successful deterrence that occurs in international politics is necessarily hidden from the event level of analysis. Consider, for example, the event series generated by successful mutual deterrence between two heavily armed countries: nothing happens on Monday, nothing happens on Tuesday, etc. This is a caricature certainly, but answers to the more serious question of how to match strategic theory to statistical data sources, and when such theories are in fact identifiable, remains underdeveloped.

Both types of objection are often reasonable, but that is not because of any inherent flaw in event data itself as a way of decomposing international politics. It can be difficult to see what else an international system might reduce to, at its lowest observational level, except a set of discrete events that are the actions and choices of agents, however those agents are understood. The problem is more about forcing events into models that were not meant for

². See <http://www.vranet.com/IDEA>