Geo-referencing Ethnic Power Relations (GeoEPR-ETH) Version 2018

Authors: Guy Schvitz, Carl Müller-Crepon
29 May 2018

GeoEPR 2018 codes the settlement patterns of politically relevant ethnic groups in independent states from 1946-2017 based on the group list in the Ethnic Power Relations dataset version 2018 (Vogt et al., 2014; see http://www.icr.ethz.ch/data). The dataset builds directly on the previous versions by Vogt et al (2014) and Wucherpfennig et al (2011).

In contrast to the previous version by Vogt et al. (2014), GeoEPR 2018 introduces four changes:

1. As in the new EPR Core dataset, GeoEPR 2018 drops the minimum population threshold of 500'000 and instead extends the coding to all independent states since 1946 according to Gleditsch and Ward (1999). As a result, GeoEPR 2018 includes new ethnic settlement data for 8 countries: Bahamas, Belize, Brunei, Comoros, Equatorial Guinea, Surinam, Tibet, Zanzibar.

2. All coding errors that were reported by users of GeoEPR 2.0 were fixed in the new version. This includes an erroneous polygon for the group 'Mizrahim (Jewish)' in Israel, an incomplete polygon for the 'Mayas' in Mexico, and several other coding errors.

3. Ethnic settlement polygons are now adjusted to country borders from the updated Cshapes 2.0 dataset, which differs from the original Cshapes data in a few instances (e.g. the border between Chad and Libya since 1946 or between Russia and Ukraine after 2014).

4. GeoEPR 2018 uses a slightly different definition of the settlement type 'Regional and Urban' (see below).

GeoEPR starts with a list of politically relevant groups from the EPR Core 2018 dataset and traces their settlement type and settlement areas over time within each country. Ethnic groups are assigned one of six different settlement types:

1. Regionally based: a group is located in a particular region/in particular regions that are easily distinguishable on a map. In line with the MAR dataset, we use the following definition: A spatially contiguous region larger than an urban area that is part of the country, in which the predominant share of group members resides. Note: Having an overlap of two groups inhabiting the same area, does not pose a problem for mapping those groups. See for example the case of Malawi, where the Yao group inhabits three different areas, some of which overlaps with other groups. In this case we make sure that we provide information on all areas inhabited by the Yao.

2. Urban: a group is located primarily urban and not in a particular region/in particular regions. We code a group as "urban" when at least 60% of the group is concentrated in cities.

3. Regional and urban: a group that falls under both the categories 'Regionally based' and 'Urban', according to the above definitions.

4. Migrant: groups with a permanent location change, e.g. nomadic groups or most Roma groups.

5. Dispersed: groups whose members are scattered throughout the territory of a state, i.e. they do not inhabit a particular region/regions but are not migrant and are always in the minority.

6. Statewide: groups that have a presence in virtually every part of the country. They do not need to constitute a majority of the population everywhere, and small pockets of territory without a significant presence can be ignored. This settlement pattern is very typical for titular nations in Europe such as the Italians in Italy, the Bulgarians in Bulgaria, or the Hungarians in Hungary.

7. Aggregate: a particular group which during a period is aggregated from several smaller ones.

Only when groups have a distinct regionally based settlement pattern, that is, when they are either coded as regionally based (1), regional and urban (3) or aggregate (6), we assign a polygon that represents their settlement area.

Statewide groups receive their country's polygon from the CShapes 2.0 dataset. For all other groups, we rely on the most adequate map material available. In a large number of cases, the GREG/Atlas Narodov Mira Dataset (Weidmann et al, 2009) was the appropriate source, either because EPR groups corresponded to GREG groups, or because various GREG groups could be treated as components of an overarching EPR group. In many other cases, it was necessary to consult external sources, such as Gordon (2005) and Levinson (1998).

In contrast to other datasets on ethnic settlement patterns, GeoEPR 2018 is dynamic, in that it records major changes in ethnic settlement patterns over time. This includes changes in international borders, changes due to mass migration and/or ethnic cleansing, as well as changes in the composition of particular ethnic groups.

The split of Yugoslavia into Slovenia, Macedonia, Croatia, Bosnia, Montenegro, and Kosovo is an example for changing state borders. We follow the CShapes 2.0 data set (Weidmann, Kuse & Gleditsch, 2010) for changes in the international system.

The expulsion of Georgians from South Ossetia in the aftermath of the Russian invasion in 2008 is an example for changing settlement patterns of an ethnic group.

Finally, groups can split into sub-groups and merge into an umbrella group. Blacks in South Africa that split into the constituent language groups after the end of Apartheid in 1994 are an example for a break-up. The Lari-Bakongo in Congo-Brazzaville that merged from two sub-groups – Lari and Bakongo – in 1969 are an example for aggregation.