Documentation for COW Colonial Contiguity Data - Version 3.1
Paul R. Hensel
16 February 2017

Overview

Version 3.1 of the Correlates of War Colonial Contiguity data identifies all contiguity relationships in the international system from 1816 through 2016 that involve colonial or dependent territories on at least one side. This includes borders between a state and a colony or dependency, as well as borders between two colonies or dependencies. Contiguity relationships involving only the territories of two states are included in the COW Direct Contiguity data set, which is available separately.

Citation

In any papers or publications that utilize this data set, users are asked to give the version number and cite the data set, as follows:

> Correlates of War Project. *Colonial Contiguity Data, 1816-2016*. Version 3.1.

Additional details of the basic coding process for the direct (interstate) contiguity data set were elaborated in the following articles, although these did not focus on the colonial contiguity data:

> Douglas M. Stinnett, Jaroslav Tir, Philip Schafer, Paul F. Diehl, and Charles Gochman (2002). "The Correlates of War Project Direct Contiguity Data, Version 3." *Conflict Management and Peace Science* 19 (2):58-66.
>
> Charles S. Gochman (1991). "Interstate Metrics: Conceptualizing, Operationalizing, and Measuring the Geographic Proximity of States since the Congress of Vienna," *International Interactions* 17 (1): 93-112.

Description of Data

The classification system for contiguity is comprised of five categories, one for land contiguity and four for water contiguity. Land contiguity is defined as direct contact between the territory of the two entities in the dyad, along either a land boundary or a river (such as the Rio Grande along the US-Mexico border). Sea contiguity is based on whether a straight line of no more than a certain distance can be drawn between a point on the border of one entity, across open water (uninterrupted by the territory of a third state), to the closest point on the territory of another entity. Four different levels of water contiguity are recorded, based on the distance between the two territories: up to 12 miles (reflecting the widely recognized 12-mile limit for territorial waters), 24 miles (reflecting the maximum distance at which two states' 12-mile territorial limits can intersect), 150 miles (from the original 1816-1965 version of the data set, reflecting what was considered the average distance that a sailing ship could travel in one day), and 400 miles (the maximum distance at which two 200-mile exclusive economic zones can intersect).

Each contiguity relationship included in this data set involves at least one colony or dependency of another state. These dependencies are listed in the colonial contiguity master data file in the "DependL" and "DependH" fields, using COW Project numeric codes. These codes are listed in the "Entities.pdf" file that is included with this download, along with the entity

names and more information about each one.

A variety of sources have been used in the process of collecting these data. The sources used for earlier versions of the data are identified in Gochman (1991); more recent updates have relied on various editions of the *Times Atlas of the World*, *National Geographic Atlas of the World*, *Hammond World Atlas*, and *Oxford Atlas of the World*.

The downloadable data archive includes a text file with this codebook, as well as three data files that are each provided in both STATA format and .csv comma-delimited format, which can easily be read into any spreadsheet or statistical software:
• **contdir.csv, dta**: The master data file (with one entry per contiguity relationship)
• **contdird.csv, dta**: A dyad-year-level summary data set
• **contdirs.csv, dta**: A state-year-level summary data set
• **Entities.pdf**: A list of colonies/dependencies that might appear in this data

Variable List
Colonial Contiguity Master File: ContCol.dta / ContCol.csv
This file contains the master records for the Colonial Contiguity data. It is the one from which all other files are ultimately derived and to which all future changes will be made. It has a non-directed format in order to prevent asymmetries from accidentally being introduced into the data set. The variable names are given in the first record in the file. The format of the remaining records is as follows:

• Dyad: A variable that combines the two states' COW country codes in a way that facilitates merging with other dyadic data. This takes the format AAABBB, where AAA is the lower value of the two country codes and BBB is the higher value.
• StateLNo: COW state number ("ccode") of lower numbered member of this dyad
• StateLAb: COW state abbreviation of lower numbered member of this dyad
• DependL: COW code for the lower numbered state's dependency (if any)
    0: No dependency for this state
• ContType: Type of contiguity relationship
    1: Separated by a land or river border
    2: Separated by 12 miles of water or less
    3: Separated by 24 miles of water or less (but more than 12 miles)
    4: Separated by 150 miles of water or less (but more than 24 miles)
    5: Separated by 400 miles of water or less (but more than 150 miles)
• StateHNo: COW state number ("ccode") of higher numbered member of this dyad
• StateHAb: COW state abbreviation of higher numbered member of this dyad
• DependH: COW code for the higher numbered state's dependency (if any)
    0: No dependency for this state
• Begin: Year and month when this contiguity relationship began (YYYYMM)
• End: Year and month when this contiguity relationship ended (YYYYMM)
• Version: Data set version number

Dyad-Year Summary Data: ContColD.dta / ContColD.csv

This file is derived from the master file and includes annual records for each pair of states that share at least one colonial contiguous relationship. To save space, the numerous dyads with no colonial borders are excluded from this file.

• Dyad: A variable that combines the two states' COW country codes in a way that facilitates merging with other dyadic data. This takes the format AAABBB, where AAA is the lower value of the two country codes and BBB is the higher value.
• StateLNo: COW state number ("ccode") of lower numbered member of this dyad
• StateLAb: COW state abbreviation of lower numbered member of this dyad
• StateHNo: COW state number ("ccode") of higher numbered member of this dyad
• StateHAb: COW state abbreviation of higher numbered member of this dyad
• Year: Year of this observation
• Land: Number of colonial land borders this year (conttype=1 in master data)
• Sea: Number of colonial sea borders this year (conttype=2-5 in master data)
• Total: Total number of colonial borders this year
• Version: Data set version number

State-Year Summary Data: ContColS.dta / ContColS.csv
This file is derived from the master file and includes annual records for each state in the COW interstate system. All states are included, whether or not they have a colonial border at any point during the period covered by the data set.

• StateNo: COW state number ("country code") of state
• StateAb: COW state abbreviation of state
• Year: Year of this observation
• Total: Total number of colonial borders this year
• Land: Number of colonial land borders this year (conttype=1 in master data)
• Sea: Number of colonial sea borders this year (conttype=2-5 in master data)
• Version: Data set version number

Version History
Version 1
The first version of the COW contiguity dataset was collected at the University of Michigan in the early 1970s.

Version 2
The first major revision was compiled by Philip Schafer of the Correlates of War project at the University of Michigan. This version updated the earlier dataset through 1993, corrected some anomalies in the previous data, and added the 400 mile contiguity relationship.

Version 3.0
For its next major revision, the dataset was split into two files: Direct Contiguity (containing data on contiguous relationships between the metropole borders of two states) and Colonial Contiguity (containing data on contiguous relationships that included at least one colony or other dependency). Paul Diehl and several graduate students at the University of Illinois updated the

Direct Contiguity dataset through 2000; several years later, Stuart Bremer and several graduate students at Penn State University updated the Colonial Contiguity dataset through 2002. At this time, the dataset was also converted to a non-directed dyad format.

Version 3.10

This was the first version of the dataset to be updated and maintained by Paul Hensel under the COW Project's data hosting program. The following changes were made in this update:

• Temporal coverage: Data extended through the end of 2016.

• Added directed dyad-year data and aggregated state-year data, neither of which was previously included with the Colonial Contiguity data set.

• Fixed over 130 entries where contiguity had been coded as continuing after one of the states left the COW interstate system (typically while one of the states in the dyad was occupied by a foreign power).

• St. Pierre & Miquelon (entity #1029) was coded in version 3.0 as becoming part of France (rather than a colony) from 1949-2000. That has now been reversed; other similar entities (e.g. French Guiana and Martinique) were still coded as French colonies during this time.

• East Timor (state #860/ETM): colonial contiguities converted to interstate upon 2002 independence (land border with #850 Indonesia, sea borders with #900 Australia and #9008 Ashmore & Cartier Islands [under Australian rule])

Future Updates

Version 3.20 will add new information to bring the Colonial Contiguity data set to feature parity with version 3.20 of the Direct Contiguity data:

• Add months and correct years: A number of entries were ended a year too early or begun a year too late so that the data didn't code the same border twice during the same year. This has been corrected in the Direct Contiguity data set by giving the month for an observation beginning or ending, which allows both phases of that relationship to be coded during the same year; 62 beginning or ending years changed with this new information. The same change will be made for Colonial Contiguity in version 3.2.

• Add notes: This data set will be expanded to add notes for every observation. These notes will explain why the contiguity relationship began/ended when it did, and (for cases of sea contiguity) will list the body of water that is crossed.

Errors and Anomalies

Every effort has been made to identify and correct errors in this data set, but some anomalies may have escaped our attention. If you detect something suspicious or questionable in this data set that the Notes field in the master data file does not clear up, please contact the data host, Paul Hensel (phensel@unt.edu). A quick response to your message cannot be guaranteed, but your input will not be ignored and will help us with the future development of this data set.