Terra Populus: Geocomputation for Population and Environment Data

David A. Haynes II [1], Tracy Kugler [1], David Van Riper [1], Steven M. Manson [2]

[1] Minnesota Population Center, 225 19th Ave S #50, Minneapolis, MN 55455,
Telephone: (612) 624-5818
Email: dahaynes@umn.edu

[2] Department of Geography, Environment, and Society, University of Minnesota, 414 Social Sciences
267 19th Avenue South, Minneapolis, MN 55455
Telephone: 612 625 4577
Email: manson@umn.edu

Abstract

Terra Populus (TerraPop) is an NSF and NIH funded cyberinfrastructure project that lowers barriers to interdisciplinary research on human-environment interactions. It has implemented a number of advances in geocomputation in order to streamline the process of making interoperable data with different formats from different scientific domains. TerraPop preserves and curates 'gold standard' data, integrating information from individual-level and aggregate census data from 238 samples in 74 countries along with global-scale raster datasets describing environmental variables such as land cover and climate. The project is developing new geocomputational methods and accompanying open-source software for integrating data from different domains and translating across data structures. We also discuss ongoing work in which we partner with other institutions to provide web-based and server-based visualization, exploratory data analysis, and modeling. The ultimate goal of these advances in geocomputation is to create a cyberinfrastructure that provides scholars and policy makers the means to discover and use data from heterogeneous sources to study the relationships between human behavior and the natural world.

1. Introduction

Terra Populus (TerraPop, www.terrapop.org) is an NSF and NIH funded cyberinfrastructure project that lowers barriers to interdisciplinary research on human-environment interactions. It has implemented a number of advances in geocomputation in order to streamline the process of making interoperable data with different formats from different scientific domains. Terra Populus (TerraPop) provides global-scale data on human population characteristics, land use, land cover, and climate change. The data are integrated across time and space, made available to researchers and the public, and preserved for future generations. The TerraPop data access system allows users to easily obtain customized datasets incorporating information from a variety of population and environmental data sources into an integrated package tailored to meet the users' needs.

2. Location-Based Integration

TerraPop preserves and curates 'gold standard' data, integrating information from individual-level and aggregate census data from 238 samples in 74 countries along with global-scale raster datasets describing environmental variables such as land cover and

climate. TerraPop leverages geographic properties to integrate information from datasets in multiple data structures to generate customized, integrated datasets for users. Users can mix and match variables originating in raster, area-level, and/or microdata datasets, and obtain output in whichever of the three structures best meets their needs. The TerraPop system uses information about the location of each record in the original data to attach contextual information, generate spatial summaries, or assign values to points, depending on the transformation requested.

3. Data Access System

The project is developing new geocomputational methods and accompanying open-source software for integrating data from different domains and translating across data structures. There is also ongoing work in which we partner with other institutions to provide web-based and server-based visualization, exploratory data analysis, and modeling. A key component of this work is development of the data access system, which offers a has a four-step process for finding and extracting data. 1) Browse Variables and Metadata: Browse variables in a variety of topics and view basic availability information; 2) Select Variables and Datasets: Many variables, especially from census data, appear in multiple datasets, so you must specify the datasets from which you want data. 3) Choose Output Structure: Depending on the variables in your cart, you can choose area-level, microdata, or raster output. 4) Define Output Options: Depending on the type of output you choose, you may need to select additional options, such as geographic level or operations to summarize raster data to administrative units. You will receive an e-mail when your customized dataset is ready to download.

4. Conclusion

The goal of TerraPop in making advances in geocomputation is to create a cyberinfrastructure that provides scholars and policy makers the means to discover and use data from heterogeneous sources to study the relationships between human behavior and the natural world.

5. Acknowledgements