

A Quantitative Analysis of the Users of Global Scale Data Sets

Karen D. Kline
PhD Candidate
Department of Geography
Assistant Secretary General
University of California
International Steering
Committee for Global
Mapping
kline@geog.ucsb.edu

With the increasing growth of spatial data infrastructures, there is an assumed increase in the amount of data made more easily available to users. There is also an assumption that users of these data sets can more easily and quickly locate, download, and use these data sets. Users are an integral part of any spatial data infrastructure, as they are the main “customer” or target. This study focuses on the users of global scale data sets to gain a better understanding of who the users are, what applications they use the data for, and what areas of the earth they study.

The information about the users of the following data sets were collected: the Global Map (Versions 0 and 1, served from Tsukuba, Japan), the Global Land Cover Characteristics (GLCC) data (served from Sioux Falls, SD, USA), UNEP-GRID Arendal (served from Arendal, Norway), and UNEP-GRID Geneva (served from Geneva, Switzerland).

Two different data sets were created from the user information. First, a comprehensive data set containing each individual user and their country, the domain (i.e., academic, commercial, etc.), the application, and the location of their research. The second was a country level data set, in which a variety of socioeconomic and internet access variables were collected, and studied in conjunction with the number of users for each country.

Several methods were used to investigate the users of global data sets. The first was the self-organizing map algorithm. This resulted in a variety of “maps” that assisted in developing hypotheses. The second was standard regression analysis, along the lines of the crosscountry or cross-national comparisons that often occur in economic or sociologic studies.

The final result is a picture of the user community that can assist in the development of infrastructure and policies at the local, regional, and global levels that can assist those users not currently able to take advantage of the data already available.

[Full paper not available.](#)