Revised Percentage of Urban Development (Version 1.1): Discussion

by David A. Hastings
National Oceanic and Atmospheric Administration
National Geophysical Data Center
325 Broadway, Boulder, Colorado 80303, USA

and James Amano
Department of Geography
University of Colorado
Boulder, Colorado 80309, USA

Background

The U. S. Navy's Fleet Numerical Oceanography Center (FNOC) in Monterey, California, created a digital terrain model in the late 1970s, by estimating maximum, minimum, and modal (most common) elevations for each 10-arc-minute grid cell superimposed on available topographic maps, generally at a scale of 1:1million. Where elevation contours were missing on available maps, the elevations were guessed at by FNOC personnel.

In addition to the three estimates for elevation, FNOC also developed estimates for the number of "ridge" lines per 10-minute grid cell, the direction of trend of these ridge lines, primary and secondary characteristics of terrain, and percentage of water cover. The data base is described in Cuming and Hawkins (1981), and presented in NGDC (1992).

The data were originally developed for internal use, NOT for archive and international distribution/reference. There is sketchy or nonexistent (or, at least, currently unavailable) documentation on the philosophy and mechanics of development of each data set in this collection. Each data set is also characterized by artifacts that affect the ability of the data to be digitally analyzed. Yet each data set is an important beginning at describing features of the land surface. The data should thus be considered a valuable (and pioneering) discussion on characterizing the land surface.

Incidentally, the version of FNOC used at NGDC is a copy received from the National Center for Atmospheric Research. This version contained artifacts that appear to result from data management problems at some stage of the process (between digitizing and final copying to NGDC), such as errors in copying disk-files. Some of these artifacts were documented by NCAR; others by NGDC (see the documentation for the FNOC data contained in NGDC, 1992). The archived version of the data obtained directly from FNOC was unreadable, apparently due to degradation of the original tape. NGDC has attempted to repair some of the artifacts in these data, though some artifacts may remain.

Spatial analysis can help us to assess the philosophy of data development; as well as find, document, and sometimes repair artifacts in the data. Such analysis can thus make the data more understandable, and more usable for many applications. This discussion provides an overview of these subjects, in an effort to help the user to adapt and use the data appropriately.

We encourage users to contribute further discussion on this data set, as well as on ALL OTHER DATA SETS IN THE GLOBAL CHANGE DATA BASE. We also encourage users to contribute proposed improvements to the data.

Artifacts in the Data

The FNOC Percentage of Urban Development data set is characterized by values over 100% in certain locations, erroneously high values (though under 100%) at high northerly latitudes, horizontal and vertical trends of "urbanization" that do not exist where shown, unlikely blocks of relatively high and low values, and other artifacts. Some of these, such as the latter two characteristics, may be the result of individual interpretations of patterns by individual data entry personnel. The blocks of relatively high and low values, for instance, may have been created by different personal biases when interpreting analog patterns on maps to digital patterns in the data set. A vertical discontinuity in Afghanistan, for example, may denote a change in digitizers. Documentation of such artifacts is not currently available.

James Amano initiated this study of the FNOC Percentage of Urban Development as part of a study project in the Geography Department, University of Colorado. The study was performed in consultation with David Hastings of NGDC. After Amano’s term of study was completed, Hastings continued with limited modifications of the data, and documentation of the effort.

Source materials for this process included the FNOCURB contained in the Global Ecosystems Database (NGDC, 1992), and Operational Navigation Charts (DMA, various dates). Editing was performed in GRASS (CERL, 1991) and a word processor.

The work consisted of the following steps:

1. The original FNOC data set contained 91 grid cells of digital value 127. These clearly erroneous values were renumbered to 255, and labeled a flag in the documentation to NGDC (1992). Inspection on Operational Navigation Charts (DMA, various dates) of all occurrences of these values showed that they occurred in areas that would normally be given a 0 value in the data set. Thus all occurrences of this value were reclassified to 0. This work was performed by James Amano, after discussion with David Hastings.

2. Northerly latitudes contained erroneously high values in unsettled or sparsely settled areas. Unsettled areas at northerly latitudes were windowed out of the whole data set, and reset to 0. A few similar values in Antarctica were similarly edited to 0. This work was performed by James Amano, after discussion with David Hastings.

3. A few unrealistically high values in South America were manually corrected. For example, an

area north of Porto Alegre, Brazil, which would have 0 value in context of its surroundings, was manually changed to 0. This work was performed by James Amano, after discussion with David Hastings.

3. In the original FNOC data set, Southern Chile and Argentina are characterized by vertical stripes of relatively low values (usually 1 and 2). These stripes were manually removed by dumping the binary file to an ASCII grid, followed by editing of this grid and reinsertion of the edited grid back into the global data set. Some other manual editing modified the sizes and distributions of cities and villages in southern Chile, Argentina, and the Falkland (Malvinas) Islands. This work was performed by David Hastings.

Table 1 shows a histogram of FNOC Percentage of Urban Development (NGDC, 1992). Table 2 shows a histogram of Version 1.1 of Percentage of Urban Development, as revised by ourselves. The two histograms were produced by IDRISI (Eastman, 1992).

Such modifications hardly make the data perfect. However, several artifacts have thus been reduced in importance, or eliminated.

How Should We Represent "Percentage of Urban Development?"

Another question also arises over the philosophical question: "What constitutes percentage of urban development?" If one city is characterized by 5 story buildings, while another city is identical but for having 50 story buildings, what is the difference in percentage of urban development? If these are different levels of urban development, what constitutes 100% development? Can anything exceed this value? If one city has larger parks, is it more or less developed? If one city has more (widely available and heavily trafficked) expressways for private cars, and another is better served by (heavily used) public transport, which is more highly developed? If two otherwise identical cities have different energy consumption, amounts of toxic emissions, or other cultural processes, how do these characteristics influence the percentage of urban development? Should agricultural or pastureland development, transportation corridors between urban centers, be reflected in this "percentage of urban development?"

Also, should there be only one data set that attempts to represent the percentage of urban development? Alternately, should there be several spatial data sets to characterize different aspects of urbanization and development, such as population density, energy consumption, alteration of the land surface, covering the land surface with impermeable materials (pavement, buildings, etc.), unnatural production/consumption of gasses and liquids, etc.? In the latter case, users could develop their own analyses, interpretations, and models of various aspects of urbanization or development, rather more sophisticated than can be covered by a single pre-defined, data set.

We feel that this topic should, some day, be addressed by the design and development of improved data on urbanization. In the meantime, this data set on Percentage of Urban Development may serve as a catalyst for further discussion (and cautious application) on this overall topic.

Conclusion

The pioneering FNOC 10-minute gridded global data base contains an initial attempt to represent the "Percentage of Urban Development." We have made some modest changes, in an attempt to repair some of the predominant artifacts in the original version of the data.

These changes may make the data more usable in certain applications. However, they do not make the data a consistent, highly analytical, representation of urbanization, development, or any related topic. A thorough philosophical discussion, design and development of several data sets as discussed just above, may be necessary before such data are available.

References

CERL, 1991. Geographic Resources Analysis Support System (GRASS) Version 4.0. Geographic information system software and documentation. U. S. Army Corps of Engineers, Construction Engineering Research Laboratory, Champaign, Illinois, USA.

Cuming, Michael J., and Barbara A. Hawkins, 1981. TERDAT: The FNOC system for terrain data extraction and processing. Technical report MII Project M-254 (Second Edition). Prepared for Fleet Numerical Oceanography Center (Monterey, California). Published by Meteorology International Inc.

DMA, various dates. Operational Navigation Charts (global coverage at 1:1,000,000 scale, revised and published periodically). Defense Mapping Agency Aerospace Center, St. Louis, Missouri, USA. (Distributed by NOAA National Ocean Service, Office of Charting and Geodetic Services, Rockville, Maryland, USA.

Eastman, J. Ronald, 1992. IDRISI, Version 4.0. Geographic information system software and documentation. Graduate School of Geography, Clark University, Worcester, Massachusetts, USA.

NGDC, 1992. Global Change Data Base, Volume 1: Global Ecosystems Data. Digital Data on CD-ROM, With Documentation. NGDC #1016-A27-001. NOAA National Geophysical Data Center, Boulder, Colorado, USA.

==

Data editing performed January-May 1993. Documentation written 11 May 1993.