Title
The Global Demography Project (95-6)

Authors
Tobler, Waldo
Deichmann, Uwe
Gottsegen, Jon
et al.

Publication Date
1995-04-01

The Global Demography Project

Waldo Tobler, Uwe Deichmann, Jon Gottsegen, and Kelly Maloy

National Center for Geographic Information and Analysis
Department of Geography
University of California
Santa Barbara CA 93106-4060
tobler@geog.ucsb.edu

Technical Report TR-95-6
April 1995

ACKNOWLEDGMENTS

Financial support by the California Space Institute, the Consortium for International Earth Science Information Network (CIESIN) through National Aeronautics and Space Administration (NASA) grant NAGW-2901, the Environmental Systems Research Institute (ESRI), and the National Center for Geographic Information and Analysis (NCGIA), is gratefully acknowledged. Substantive support has also been received from many cooperative individuals and from the following agencies, generally in the form of data:

Australian Bureau of Statistics, Canberra
Birkbeck College Department of Geography, London
Centro Internacional de Agricultura Tropical, Cali
Demosphere International, Fairfax
International Food Policy Research Institute, Washington D.C.
International Livestock Research Institute, Nairobi
National Geographic Society, Washington D.C.
New World Demographics, Shady Side
Oak Ridge National Laboratory, Oak Ridge
Rijksinstituut voor Volksgezondheid en Milieuhygiene, Bilthoven
Statistical Office of the European Union, GISCO, Luxembourg
Statistics Canada, Geography Division, Ottawa
Statistics New Zealand, Wellington
United Nations Environment Programme, Global Resource Information Database (GRID),
    Arendal, Sioux Falls, Tsukuba, Geneva, Nairobi
United Nations Food and Agricultural Organization, Rome
United Nations Statistical Office, POPMAP project, New York
U. S. G. S. EROS Data Center, Famine Early Warning System, Sioux Falls
World Resources Institute, Washington D.C.

The opinions, conclusions, and recommendations are those of the authors and do not necessarily represent the view of any of the sponsoring organizations.

ABSTRACT

Demographic information is usually provided on a national basis. But we know that countries are ephemeral phenomena. As an alternate scheme one might use ecological zones rather than nation states. But there is no agreement as to what these zones should be. By way of contrast global environmental studies using satellites as collection devices yield results indexed by latitude and longitude. Thus it makes sense to assemble the terrestrial arrangement of people in a compatible manner. This alternative is explored here, using latitude/longitude quadrilaterals as bins for population information. This data format also has considerable advantage for analytical studies. The report is in three parts. Part I gives the motivation and several possible approaches. Ways of achieving the objective include, among others, simple centroid sorts, interpolation, or gridding of polygons. In Part 11 the results to date of putting world boundary coordinates together with estimates of the number of people is described. The estimated 1994 population of two hundred seventeen countries, subdivided into nineteen thousand thirty two polygons, have been assigned to five minute by five minute quadrilaterals covering the world. The grid extends from latitude fifty seven degrees south to seventy two degrees north latitude, and covers three hundred and sixty degrees of longitude. Just under thirty one percent of the (1548 by 4320) grid cells are populated. The number of people in these countries is estimated to be five billion six hundred eighteen million, spread over one hundred thirty two million square kilometers of land. Part III describes needed extensions, and the appendices contain detailed information on our results with maps and data sources.

PART 1: GLOBAL DEMOGRAPHY

The Global Demography Project has as one of its objectives the assembly of world demographic data on the basis of latitude - longitude quadrilaterals. These data are to be made available on machine readable media. We report first on why this should be done, how it can be done, and what can be done with the results.

WHY

Global change studies have predominantly concentrated on studies of physical phenomena which describe such issues as climatic change or the depletion of the ozone layer. Many of these problems are human induced or affect human activities. Only recently has interest shifted towards intensive study of these human dimensions of global change. One reason for the small number of studies which incorporate human factors in global change modeling is the lack of suitable population related data. In general terms these data encompass demographic variables (population numbers, components, and dynamics), spatial arrangement, social factors (such as attitudes, education, income, health, etc.), and variables relating to the economic activities of people. On the national level many of these variables are collected and compiled on a regular basis by various agencies and institutions (e.g., the World Bank, World Resources Institute, the United Nations, etc.), and macro level economic and social analysis has been made easier by the increasing availability of these data in digital form either as a generic database or in a geographical information system (GIS) format.

On the subnational level, however, efforts at compiling socioeconomic data have been scarce. Indeed, to date no consistent global data sets of socioeconomic indicators for subnational administrative units exist. Partly this is due to the fact that data collection is rarely coordinated between countries. While great efforts have been made for national level data - for example, the national systems of product and income accounts designed by the U.N. Industrial Development Organization (UNIDO), or the standardized medical reporting encouraged by the World Health Organization (WHO) - there are no generic reporting schemes for the same information on a subnational scale. A second factor is that the collection of socioeconomic variables is a process which is often far more tedious than the collection of physical data. For instance, large scale monitoring techniques such as remote sensing have no equivalent in the human domain. Data collection therefore has to rely on expensive and time consuming censuses. For a comprehensive overview of these problems see Clarke and Rhind (1992).

The global change research community is one of the most important groups interested in global information relating to socioeconomic variables. And this implies that these data need to be rendered compatible with those commonly used in global scale analysis of physical phenomena. In most instances, such factors, e.g. vegetation cover or climatic variables, are stored in gridded form. An interesting aspect of this is to attempt a description of the global distribution of population related factors via a system of mathematical equations in a form compatible to those used by geophysicists Jobler 1992).

The focus of global change studies is also often to link environmental data to socioeconomic variables for the study of human induced degradation processes. A complement to this is the study of the social impacts of environmental degradation. Examples are the increasing problem of environmental refugees, or, as a specific example, the study of the relationship of malnutrition and deforestation in Zaire by Fowler and Barnes (1992).

The range of potential applications for global, socioeconomic data is clearly very wide. The lack of consistent population related information was seen as one of the major impediments for strategic planning and targeting of development strategies for large areas at the recent meeting of the Coordinative Group for International Agricultural Research (CGIAR) and Global Resource Information Database of the U.N. Environment Programme (UNEP/GRID) on digital data requirements in the agricultural research community (Martin 1992). Socioeconomic data on a regional, or a continental, level are used by the CGIAR centers for the selection of representative survey sample sites, for the generalization of relationships determined in micro level analysis to larger areas and for the forecasting of future target areas. These efforts are all based on the relation of population pressure to food availability. Further application can be anticipated in global level migration studies, large scale epidemiological research (e.g., Thomas 1990 and 1992), or global economic modeling (Hickman 1983).

Many of the human impacts on the natural environment are directly related to the number and arrangement of people, and the associated industrial metabolism. This project attempts to address a small part of this problem by collecting population totals at a modest level of resolution, but for the entire earth. There are two primary advantages to having global demographic data referenced to latitude and longitude. The first of these advantages is the analytical capability which it provides. The approach renders geographic observations susceptible to the same kinds of objective analyses as are performed on time series data, as pointed out by the Swedish geographer T. Hagerstrand in 1955. Viewing spatial data as a kind of (two dimensional) regional series, we may undertake a variety of analyses Jobler 1973). For example, population pressure might be defined analytically as the absolute value of the gradient of population density, and is easily computed from gridded data (Muehrcke 1966). Thus analytical and simulation models of several types are more conveniently formulated when one works with cellular geography (Tobler 1979b). Administrative partitionings of a country or region are generally irrelevant to scientific work. More recently a pilot study has been conducted of some practical advantages of such data reporting (Clarke & Rhind 1992), including the potential impact of sea level rise on inhabitants of the coastal region of a Scandinavian country.

The second motivation, as outlined above, is in the ability to relate demographic data to other global data. Global monitoring systems using satellite sensors report data by pixels, small, usually square, pieces of the terrestrial surface. Population data typically are available only for political units, generally countries and their political subdivisions. There is thus a mismatch between the two spatial data collection and reporting schemes. The higher the resolution of the satellite system the greater the mismatch. This discrepancy renders more difficult the inclusion of population information into global modeling efforts. Minor additional advantage accrues to the spherical quadrilaterals because these are essentially invariant collection boxes. They have a permanence which does not apply to political units, which are likely to change. A corresponding disadvantage is that political action is generally only available on a nation state or regional basis. There have also been suggestions that ecological zones are more appropriate units of study than countries or other political units. Unfortunately there does not seem to be agreement on which set of ecological units to use. Such units are also likely to change over time, though probably not as rapidly as political units. Having the information in small lat/lon boxes naturally also allows reaggregation to arbitrary partionings of the earth's surface, including ecologic zones. Another alternative, not used, is to map the entire surface of the earth into an equal area square via a map projection and then to use this as the basis of a quadtree structure Jobler & Chen 1986). Watersheds, or a hierarchy of watersheds, have also been proposed as a useful way of organizing information about the earth.

The idea of using the earth grid to assemble data is not particularly new. In 1851 Lt. Maury published a map of whale sightings by five degree quadrilaterals. Of course the oceans were not then partitioned politically, so he had to use this method. Amusingly the movie "Moby Dick" uses a map based on such observational recording. The book does not contain this scene, although a footnote refers to Maury's work. In Sweden population counts in 1855, 1917, and 1965 were assembled in 10 kin by 10 km cells. Since the introduction of computerized geographical storage and processing systems in the 1960's the process has accelerated. Many of these systems work on a raster or grid basis and much local data is now collected in this manner. Minnesota, for example, has an information system which uses the square-like Public Land Survey system of Sections, Townships and Ranges for many observations. A recent British Census reported data by kilometer squares on their national grid (Rhind et al 1980). The recent Population Atlas of China (Chengrui et al 1987) contains several maps with population density shown within 1'15" by 1'52.5" quadrilaterals. Japanese social data are now sold, for marketing studies, by a commercial firm on the basis of 500 meter by 500 meter cells for a portion of the country.

We must also address the question of which demographic data. The typical demographer wishes population totals, further broken down by gender, age, occupation, education, and so on. Naturally the number of births, by age of mother, and the number of deaths, by age groups, should also be known. Completing the demographic equation requires knowledge of the number of inmigrants and of outmigrants, optimally also by age group and social, occupational, and educational status. In the migration category refugees should also be included. The economist would want information on energy consumption, capital investment, and so on. The political scientist is interested in attitudes. The geographer would want tables of trade or of information exchange. The list seems endless. Of course all of the information should pertain to the same time period. Our effort represents only a very modest beginning and one of the concerns is the necessity to introduce mechanisms by which other groups, international, national, and/or private, can be induced to continue to maintain and improve the database established. In order that this be done it is necessary to overcome the reluctance of many nations to release information that may be embarrassing or considered of value in potential conflict, including economic domination. And it is important to emphasize that the actual indirect impacts of population are even more difficult to assess, with even greater data lacunae, than simple population quantities (Stem et al 1991; National Research Council 1994). Sociopolitical data and attitudes may be as important as socioeconomic or demographic information.

The United Nations publishes a considerable quantity of statistics, and there are now several sources of machine readable world population data. The number of countries in the world ranges from 165 to 311, depending on the source and is partly a matter of definition (Freedman 1993). Besides the United Nations and some other official (i.e., bureaucratic, governmental) public agencies, there are also several private (e.g., atlas publishers) and non-profit (Population Reference Bureau, etc.) organizations that have assembled such data. Generally they do this only by political unit or subunit; in some cases by city. In addition there are several publications that focus on global problems sponsored by the World Resources Institute, by the Worldwatch Institute, by the World Bank, and by several others. These provide tabular data by country, and cite sources, including national and international documents. Recent books include Keyfitz & Flieger (1990), Willett (1986), World Resources Institute (1990), UNDP (1990), and Zachariah and Vu (1988). A missing component in virtually all of these information assemblages is the geography, i.e., the precise specification of the coordinates of the boundaries of the units used. This is where geographers and geographical information and analysis systems can provide valuable assistance. Furthermore the quality of the data given by these suppliers of data is extremely heterogeneous and of unknown validity. Internal strife often renders all census data inadequate. One need only recall what is happening in Rwanda, Yugoslavia, the former Soviet Union, Somalia, the Sudan, and Angola to see the difficulty. On occasion political considerations bias "official" statistics to reflect favorably on the group in authority, or to justify action. At a later stage in this type of work a necessary component should be studies to evaluate the problem of data accuracy. Sub-national data are usually the most difficult to obtain at a distance, and often may be less reliable.

HOW

Several methods exist to make compatible the disparate data bases. One can aggregate the environmental pixel data to the country or regional scale, which has the advantage that decisions are often made at this level, but this also often results in meaningless averages. Consider, as an example, one average temperature reading for the entire United States! Mixing the Alaskan climate with that of Florida hardly makes sense. Alternatively one can attempt to reallocate the national data to finer levels of resolution. This is what is described here, and has involved correspondence with many international and national agencies. We attempted to assemble global population data (neglecting related parameters) as counts and to convert them to spherical quadrilaterals, small latitude and longitude cells. Using one degree as the bins we would get 180 x 360 such cells. If only land areas are included, about sixteen thousand cells, each of approximately 7,800 square kilometers. This contrasts with about 185 countries, with an average size of over 700,000 square kilometers. The resolution of these two schemes, the one degree cells or the countries, is very coarse for global modeling purposes. Finer data are generally available for political subdivisions of countries, although difficult to obtain outside of the individual nations. In the United States there are 50 states, 3,141 counties, and some 40,000 enumeration areas. Data are even available at the city block level. In France there are data for over 32,000 communes, in Switzerland there are 26 Cantons and 3,029 Gemeinde, India has 337 districts, Mexico over 2,000 Municipio, Italy records information on about 60 million people in 8090 municipalities, etc. The resolution also varies within different parts of each country. In the United States the counties increase in size as one moves westward and across less densely occupied territory. The reason for this is a mixture of history, physical terrain, climate, and technology. But if one received a piece of film from a manufacturer with resolution that varied from one part to the next by this order of magnitude one would certainly reject it as being unusable. Have you ever heard of a variable resolution film or camera? They may exist. Of course the resolution used by data collection agencies throughout the world to some extent reflects their perception of national problems, or population density, and is a type of adaptive grid. Some partial differential equation solvers use variable resolution grids, and these would seem to be possibilities for global modeling too. Forcing everything into a fixed grid carries some obvious disadvantages. But now it is time for some definitions.

For the purposes of the present study we define several levels of political partitioning. The nation state or country we refer to as the zero level. Then comes the first level, comparable in the United States hierarchy to states. Below this is the second level, equivalent to US counties. In many countries, including the United States, there are further levels. We have attempted to obtain data and boundary coordinates to only the second level, but have not been successful for the entire world. The second level for a country such as Switzerland consists of 3,029 Gemeinde, very nearly the same as the 3,142 units available for the United States. Switzerland is a country of approximately 41,150 km^2 to the USA's 9,282,840 km^2. The same number of units at the second level clearly yields a different level of spatial resolution in different parts of the world. We define the mean spatial resolution in kilometers by the following formula

$$\text{Average Resolution} = \sqrt{\frac{Area\ \ (km^2)}{No.\ \ of\ \ observations}}$$

Here the number of observations is the number of data collection units. Resolution measured in this manner is clearly the same as the length of the side of a square whose area is that of the average observation unit. A useful additional parameter would be the variance of this quantity for each set of units for which the resolution is calculated.

For the sake of comparison the resolution (in meters) of a map is conveniently calculated by dividing the map scale by 2000. From the Nyquist sampling theorem we know that a pattern can only be detected if it is greater in size than twice the resolution. This assumes data without error. When there is error in the data (is there ever no error?) then patterns or features can only be detected if they are several times this size. Using map scale again as a convenient analogy, the optimistically sized pattern that can be detected on a map is, in meters, the map scale divided by 1000. Thus on a map at a scale of one to a million (1: 1,000,000) objects greater than one kilometer in size might be detected. We can now make some comparisons between spherical quadrilaterals and map scale.

TABLE I

Spherical quadrilaterals	Number of cells	Average size	Equivalent map scale
5° x 5°	36 x 72 or 2592	~197,000 km^2	(very small)
1° x 1°	180 x 360 or 64,800	7,900 km^2	1/178,000,000
5' x 5'	2160 x 4320=9.3 x 10^6	54.7 km^2	1/15,000,000
1' x 1'	10800x21600=2.33 x 10^8	2.2 km^2	1/3,000,000
20" x 20"	(large) ~ 2.10 x 10^9	.25 km^2	1/1,000,000
5" x 5"	(large) ~3.36 x 10^{10}	15,000 m^2	1/250,000
180 countries	(land area only)	453,392 km^2	(very small)

The numbers in the table are rough approximations but suggest the magnitude of the difficulty. The ultimate resolution needed for global population modeling is problem dependent but, using Table 1, it is seen that having data by one degree quadrilaterals of latitude and longitude is the equivalent of working with maps at an average scale of approximately 1/178,000,000. The mean quadrilateral size is then 7,900 square kilometers. A sixty fold reduction would result in quadrilaterals of one square minute of arc, roughly one and a half kilometers on a side. About thirty seven million cells of this resolution size would capture the world population, taking 3/4 of the earth's surface as unoccupied water with Antarctica making up an additional nine percent of the land area. This is a large amount of data and corresponds to a global map scale of approximately one to three million (1/3,000,000). A more ambitious effort would attempt 20" by 20" quadrilaterals, finer than the largest scale (1/1,000,000) world map for which a digital version now exists (Danko 1990). A map coverage of the entire world at a scale of 1/250,000 would correspond to a demographic data base of 3.4 trillion 5 " by 5 " quadrilaterals, with cells averaging roughly 125 meters on a side. This is obviously impractical at present, and, since human populations are mobile, hardly useful in any event. The average daily activity space of individuals is culture, environment, social, and urban-rural status dependent, but averages more than fifteen kilometers in western societies. This argues strongly for a macrogeographic approach, along the lines pioneered by W. Warntz (1965). In this paradigm populations are modulated by their distance apart, to yield a "potential" field covering the earth. It seems reasonable that population potentials based on raw counts of people, or on births, deaths, migrations, or on income weighted, or on energy-usage weighted populations could usefully be related to global models. The potential approach is particularly useful when events depend on many neighboring places so that place specific data capture only part of an impact. The population potentials (computed at each place as the sum of a geographically discounted numerical value) are more easily calculated from population data given by spherical quadrilaterals. Geodemographic models of the accounting type - i.e., future population at every geographic location equals current population plus births less deaths plus immigration less emigration - are also more easily modelled as a system of partial differential equations (Dorigo and Tobler 1983) when each of these components is available in a system of regular tessellations rather than by political units. This approach, detailed in Part III of this report, essentially makes use of the finite difference form of these equations, a relatively easy task on modem high speed computers.

At this time it is not possible to specify the optimum resolution for global demography. Tables and maps in the appendices indicate what we have achieved. Of course the resolution might have to vary in different parts of the world and for different purposes. At present world population data only exist by political collection units. Our method has been to examine these data from a variety of sources. We have then attempted to convert the population figures by one of the several possible methods detailed below. The basic data are the numbers of people, but the components of change (births, deaths, immigration, emigration), age distribution, and energy consumption are also important; several of these are already available as estimates by country (or subunits). There are of course many problems with these data, synchronicity being the most obvious, and one might wish a time series for each latitude/longitude cell in order to do space-time modeling and animations (Tobler 1970).

Geographical locations can be specified in many ways. The different names given to the same place are referred to as its aliases. The processing of geographical information often requires conversion between these aliases. There are many examples. In an academic environment latitude and longitude are often taught as an appropriate way of naming places. But the US postal service does not use this system, though it could, as is evidenced by the accompanying illustration showing a post card which arrived through the mail (Gould 1990). A few of the other geographic conversions which occur in practice are given in the next table (Table 11). Conversion of information given by political unit to a grid of latitude/longitude cells is only one such conversion task (Tobler 1990a). Eventually one would expect nearly all of these conversions to be available on demand, using as input any data source with the output format tailored to the expected usage. It is also desirable to develop analysis methods which are invariant under alternative locational naming conventions (Tobler 1990b).

A simple method of converting national data to spherical quadrilaterals is to assign the average national density to all of the lat/lon cells covering the country, with simple prorating for partial cells. This yields a piecewise continuous function for the world, with abrupt jumps at the edges of the countries, and a constant value within each country or unit. This method is clearly not appropriate for large countries, but must be used for such aggregates as gross national product (GNP), by definition only applying to the political unit, and a context variable for other properties. An alternate method of approximation assigns all of the data to the centroid of the country (or subunit) polygon. One then simply adds all of the values for those centroids that fall into each cell, a rather simple binning procedure. For this all that is needed is the centroid of each collection unit. This technique was used by Haaland & Heath (1974) with enumeration districts for the United States and by Tobler (1992) for the entire world using five degrees of latitude and longitude. In this method there may result some cells with no population, and others with excessively large values. A light smoothing of this crude assignment helps to redistribute the population.

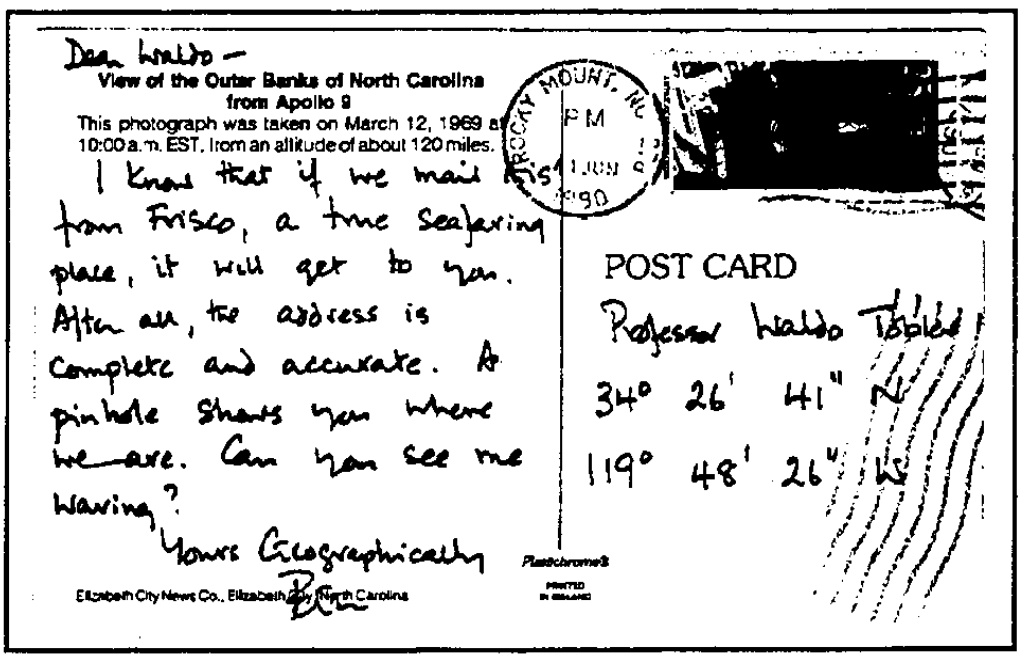


TABLE II:

A FEW EXAMPLES OF GEOGRAPHIC TRANSFORMATIONS

Not exhaustive and oversimplifies; also note the difference between > and < >

POINT > POINT
(All aliases for the same place)
Street address < > State plane coordinates
Lat/Lon < > U.T.M. coordinates
Digitizer x,y > Lat/Lon
Rectangular x,y < > Polar (distance, direction)

POINT > LINE
Highway Accident < > Highway segment

POINT > AREA
Street address > Census tract
Place name > Bingo coordinates
Lat/Lon < > Public Land Survey

POINT > FIELD
Spot elevation < > Contour map
Dot map < > Density contours

AREA > POINT
Census tract < > Centroid
Public Land Survey < > Lat/Lon

AREA > LINE
Census tract > Street names
Political district < > Address list

AREA > AREA
Census tract < > School district
Country < > Grid cell
Political Unit < > Ecologic zone

AREA > FIELD
County population < > Contour map
Census data > Choropleth map

FIELD > POINT
Density contours < > Dot map
Elevation contours < > Spot height

FIELD > AREA
Contours > Polygon total

FIELD > FIELD
Contour map < > Gradient map
Contour map < > Smoothed contours

Another alternative is to interpolate from the known observations to the spherical quadrilaterals. Given the densities at centroids one can interpolate from these to a lattice or triangulation , making the usual assumptions needed in interpolation from data given at points. This is the method used by Bracken & Martin (1986) and Martin & Bracken (1991) for areas in Great Britain. A better method is to use smooth pycnophylactic redistribution to reassign populations from one arbitrary set of regions to another (Tobler 1979a; Rylander 1986; Park 1990). This reassignment method has the advantage of always correctly summing to regional totals, uses regional boundaries rather than centroids, and takes into account values in geographically adjacent areas. Coordinates for the boundaries of political units are more difficult to obtain than coordinates for centroids, but many such boundaries, at the national level, are now available in computerized map drawing systems (e.g., Taylor 1989). The computational load for this method, involving a finite difference approximation to the solution of a partial differential equation, is greater than the summation methods previously cited in this report, although the accuracy of the resulting interpolation should be better. Since this method has actually been used in this project it is summarized in the adjacent insert. An appendix gives details on how to adapt the published version (Tobler 1979a) to take into account the shape of the earth, assumed spherical. In this approach it is not reasonable to use demographic rates; one must use counts, e.g., of people, births, deaths, migration, or diseases, etc. This also holds true for the ratio known as population density since, as with all such rates, the value obtained is highly dependent on the spatial extent of the chosen denominator. Rates can later be obtained by division of one cellular map (i.e., computer grid array) by another.

One can also make measurements on population density maps, where these are available (c.f., Rhind et al 1980; Liu 1987). Using this method requires digitization of the isopleths followed by interpolation of the density values from the isopleths to a grid, then the use of a mathematical integration to get the "volume" contained within each grid cell, taking into account the map projection used. This is rather like the "cut and fill" computation used by the civil engineer when designing road construction. Urban populations on these maps are often represented by circles whose area is related to the population, and this can be inverted if class intervals (a form of discretization) have not been used. Population density maps are usually at a small scale so that this technique can only result in a data file at a coarse resolution.

It is also well known that it is possible to make population estimates from aerial photographs and satellite images, though the accuracy of such estimates is disputed (Dureau, et al 1989). The procedures are well established for land use and agricultural crop estimates but also work for nucleated settlements (Tobler 1969b; Lee 1989). Such imagery can also be used to mask empty districts. The method is best suited to high resolution estimates covering small areas, and is not well suited for complete world coverage in a short time. And no one argues that birth and death rates, and similar important demographic components can be obtained in this manner.

MASS PRESERVING REALLOCATION FROM AREAL DATA.

First define the primary condition for mass preservation. This is the required invertibility condition for any method of areal information redistribution:

$$\iint_{R_i} f(x,y)\ dx\ dy = V_i \quad \text{for all i,}$$

where V_i denotes the value (population in the present context) in region R_i (a subnational polygon).

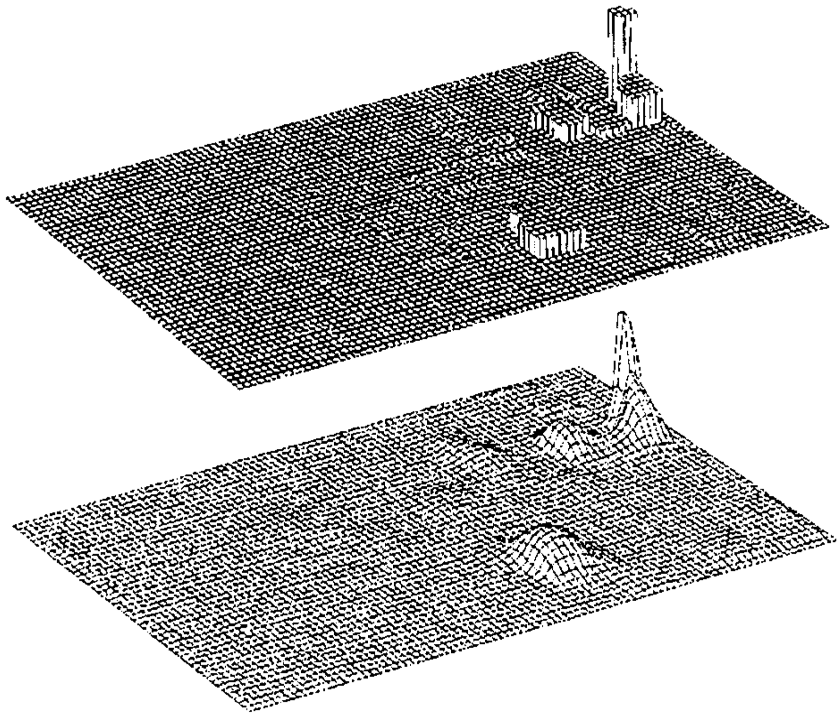
Next constrain the resulting surface to be smooth by requiring neighboring places to have similar values. This is an assumption about spatial demographic processes, a form of geographic insight capturing the notion that most people are gregarious and congregate. Densities in neighboring areas therefore tend to resemble each other unless there is a physical barrier. Laplacian smoothness, the simplest kind, is obtained by minimizing:

$$\iint_{R} \left(\frac{\partial f}{\partial x}^2 + \frac{\partial f}{\partial y}^2 \right) dx\ dy$$

where R is the set of all regions. The boundary condition generally used is

$$\frac{\partial f}{\partial \eta} = 0 \ .$$

This says that the gradient normal to the edge of the region is flat; it is one of several possible choices. More detail is given in the original paper (Tobler 1979a). Also see the appendix to this report. The example below illustrates the procedure graphically for 1990 population by county for Kansas.



PART II: WORLD POPULATION AT MEDIUM RESOLUTION

As a first step towards a global socio- political- demographic- economic- database indexed by latitude and longitude we have produced a consistent data set of population numbers for the world. This required the collection of two major components. The first component was spatial boundary information describing administrative regions within the countries of the world, followed by assembly of the corresponding enumerations (or estimates) of the total population residing in each of these regions. In this initial attempt only information on the total number of people was collected. From these two components we have derived a gridded data set of population numbers for the globe at a modest resolution. We refer to this as modest or medium resolution because, as described earlier, finer data exists for several parts of the world. But we are here not concerned with data at the city block level. And we recognize that simply knowing the geographic arrangement of the world's population is not sufficient to determine its two-way interaction with the environment.

Information on 217 countries (or units; included are all major countries, but also small islands such as Aruba and Guam) has been obtained from various agencies and institutions. In some cases we have been able to assemble second level administrative unit boundaries but not second level populations. In other cases the converse situation holds. The exact details are in the appended large tables. The primary product, an approximation to a globally continuous surface of population, required overlaying a grid on the polygons, followed by an assignment of the population to the cells of this grid. Essentially the problem is to distribute the published population count for a given administrative unit over the grid cells that fall within the unit, taking into account the population values in the adjacent units.

BOUNDARY DATA

The key to producing a spatial demographic database are the geographic boundaries that are used to reference socioeconomic information. Virtually every country of the world has instituted an administrative hierarchy that partitions the country into units; often an attempt is made to have these be more or less homogeneous units. For the purposes of this project we define the country to be the zero-level within the hierarchy. In the United States, the 50 states (plus Washington D.C.) define the first subnational level, the 3141 counties represent the second level, and so on. The size of these units obviously varies greatly by country. Most first level units in Zaire, for example, are still larger than the neighboring countries of Rwanda and Burundi. For that reason, objective measures such as the average size of a unit at a given level (i.e., mean resolution as previously defined) or the number of people per unit in a country provide a basis for cross-national comparison. These measures are generally more appropriate than the position or level of the unit in the political hierarchy.

We attempted to obtain second level boundaries, corresponding to counties or districts, except in the case of very small countries or islands. We refer to this as medium resolution because this level of detail bridges the gap between the coarse resolution of national level data and high resolution at the block or enumeration area level. Due to the limited time and resources available, the global demography project relied mostly on existing spatial boundary data sources. The primary producers of such data are national or supra-national public data collecting agencies (e.g., US Census Bureau, Eurostat), international institutions (e.g., UN-FAO, UNEP, CGIAR), universities engaged in research activities around the world, and private companies providing spatially referenced socioeconomic data chiefly for marketing applications. For a few countries for which no existing boundary data could be obtained, administrative boundaries were digitized at NCGIA using maps found in the University of California system libraries. We have included a detailed list of data sources by region in the database documentation (see appendix).

The global database assembled at NCGIA contains over 19,000 administrative units for some 217 countries of the World. This count includes several small separate islands. The largest number of units were available for the United States (3141), China (2422), Mexico (2402), and Brazil (2224). For very small countries, including some of the island states of the Caribbean or the South Pacific, no attempt was made to obtain subnational boundaries. Since the boundary information came from a wide range of sources, considerable effort had to be spent on rendering these data compatible in order to create a consistent global database. This typically involved coordinate transformation and projection change as the first step. We chose geographic (latitude & longitude) coordinates in decimal degrees for this project because this is the most generic reference system for spatial data. For some data sets, conversion to ArC/Info format was required prior to the coordinate conversion. We used Arc/Info running on the NCGIA IBM RS/6000 workstations as the platform for data compilation, editing and storage. Individual coverages for countries or groups of countries were merged into one of several regional coverages. The maps presented in the appendix show the boundaries of the administrative units used for the individual regions. These maps also give a good indication of the accuracy of the interpolation since the redistribution of the total population within a unit will be more accurate if the units are very small.

In order to merge the individual regional coverages, the international boundaries needed to be matched to avoid sliver polygons or polygon overlaps caused by differences in data sources. These regional coverages formed the basis for setting up the population data and running the gridding routines. While this whole process is conceptually straightforward several problems had to be dealt with, including:

- Many of the spatial data sets obtained had no documentation, or only minimal documentation, as is apparent from the database descriptions. Thus information on the scale and reference of the source maps, the time period to which the boundaries refer, or the accuracy of the information was often unavailable. For a number of sets of data, even the projection of the maps was not known. Administrative boundaries are rarely available on maps with precise geodetic referencing. Particularly in developing countries one needs to work with whatever is available, including “cartoon maps" which are often page size, hand-drawn sketch maps provided in census publications. Such information may be better than no information at all and may indeed reflect the fact that at this level administrative boundaries are often not well defined. In cases where the boundaries were not spatially referenced, rubber sheeting using standard Arc/Info tools was required to make the coverages compatible with the global database. The area of the polygonal units was obtained by converting the latitude and longitude coordinates of the boundary points to a local equal area map projection (by Carl Mollweide) with a subsequent polygon area computing routine. This introduces an error which is small as long as the units are small. Summing these values for all of the polygons yields an estimated 132,306,314 square kilometers of occupied land, 25.94% of the world's surface area. The areal estimates may also differ from published figures because some countries include or exclude internal water bodies, and have differing definitions of how large a lake must be to be excluded. Despite these source limitations we feel that the database is sufficiently accurate for global or regional scale applications, but not for high resolution applications which require more detailed boundary information.

- The source scale and the resolution at which administrative boundaries were digitized varied a great deal. Our estimate is that the boundary data sources ranged in scale from about 1:25,000 to about 1:5,000,000. This is an estimate because in most cases scale information was not indicated on the source maps.

- Due to the limited time available, no effort was made to make the international boundaries compatible with "standard" global country databases such as the Digital Chart of the World at 1: 1 million scale or the World Boundary Database II (WBDII) at 1:3 million scale. In practice two neighboring countries were merged into the same coverage and the resulting sliver polygons were eliminated using standard Arc/Info tools. In retrospect this was a mistake and in the future it would be desirable to standardize the database by replacing all international boundaries with a standard data set. In this manner it would be possible to avoid problems encountered in combining nations or regions. An international effort may be required to achieve this. We have ignored problems of *de-jure* versus *de facto* boundaries and of disputed boundaries.

- The spatial resolution of the administrative units acquired for different countries varies greatly. We normally used the most detailed level available to us. For a few countries more detail is available (e.g., block level data for the United States), but the data volume involved then becomes impractically large. The average resolution by country is included in the appended tables and on the CENSUS.LST file on the PC diskette. Also listed there is the range in census unit sizes for each country, giving an impression of the resolution ranges. Taking into account the total occupied land area and the number of units, we calculate an average resolution for the world land area of 83.4 kilometers, which is directly comparable to the spherical quadrilaterals used (see TableIII).

- Boundary data for several countries (e.g., Australia, Canada, Community of Independent States, Mexico, New Zealand) come from commercial sources or from public agencies that put restrictions on further distribution of the data. This boundary information is copyrighted and consequently cannot be distributed freely. The database documentation lists all sources and contact addresses for these data sets. Interested parties can contact these organizations directly.

POPULATION DATA

For each of the administrative units we attempted to obtain a recent estimate of the resident number of people. We of course recognize that simply knowing the geographic arrangement of the world's population is not sufficient to determine its two-way interaction with the environment. For many analytical purposes it would be useful to have other demographic indicators as well, e.g., population by age and sex, rural versus urban, and economic involvement, or time series data of population figures. The compilation of such data, in particular the effort involved in ensuring cross-national definitional compatibility, was beyond the scope of the current project. In several cases the agencies from whom population and/or boundary data were obtained do also have these sorts of additional demographic or economic data.

Population figures were typically taken from the latest available census. A list of censuses by country is compiled by the UN Statistical Division and updated continuously. Additional sources of demographic data are civil registration systems or estimates based on sample surveys. For this project we used the most recent enumerated or estimated population figure released by the national statistical offices, and published in official census publications, year books, or gazetteers, that we were able to locate given our resource limitations.

For most countries, data were available from the 1980s or early 1990s. Often there is a considerable time lag between a census and public dissemination of the data. Thus the most recently collected figures may not yet have been published. All limitations associated with published demographic data, which are well known in demographic research (see, for example, Chapter 3 of UNFPA 1993), obviously apply to our collection of population figures as well. Regrettably, the sophistication of mathematical models conceived by demographers stands in stark contrast to the uncertainty associated with the population data that demographers have available. The main problem faced in compiling the count data relate to synchronicity. Firstly, census dates vary. In our database, the reference years range from 1979 to 1994. We decided to derive estimates for 1994 for administrative units using a standard growth model based on average annual growth rates by country provided by the United Nations (UN 1993, table A.2):

$$P_t = P_c \exp\left(\sum_{i=c}^{t-1} \frac{r_i}{100}\right),$$

(see Rogers 1985) where Pt and Pc are the regional populations in the target year (1994) and in the census year respectively, and ri is the annual percentage growth rate in year i for the particular country. While this method is admittedly simplistic (it assumes uniform growth in all regions of a country) the error introduced should be small for most countries that have had a recent census. Furthermore, the error is likely to be minor in relation to the overall uncertainty associated with census information. Ideally we would like to use historical census data to derive growth rates that are specific to each administrative unit in order to capture recent urbanization processes, migration, etc. This was not possible with our limited resources. We have however compared our estimates with the latest UN estimates and the disparity between these two suggests the magnitude of the possible errors. The details are in the tables in the appendix. As another comparison the 1995 World Almanac on pages 840-841 lists 1994 estimates of the population of the major countries of the world. These numbers, which differ from our two lists, were allegedly obtained from the US Bureau of the Census. The total 1994 population estimated for our 19,032 units is 5,617,519,139 people. This works out to a world average of 42.5 persons per square kilometer of occupied land.

The second problem related to synchronicity derives from the fact that boundary and population data often did not come from the same source. Administrative boundaries change frequently, and in some cases, boundaries were only available for a previous census, and did not match the units for which population figures were defined. In other cases, the opposite was true. In the few cases where a mismatch occurred, the population data had to be aggregated or disaggregated using simple proportional areal weighting, although more complex areal interpolation techniques could be applied (e.g., Goodchild et al 1993). Attempting to compile geographical time series exacerbates the problem of boundary consistency and it may be necessary to use pycnophylactic areal redistribution between changed zonations to make the data compatible over time.

GRIDS

The primary product, an approximation to a globally continuous surface of population, required the gridding of the polygons followed by an assignment of the population to the cells of this grid. Essentially the problem is to convert the published population count for a given administrative unit into a population within the grid cells that cover the unit. This grid assignment is followed by a smoothing to overcome the jumps between boundaries of the administrative units by taking into account the population values in the adjacent units. The smoothing technique used is the smooth pycnophylactic reallocation described earlier in this report, modified to take into account the spherical shape of the planet. The routine has been implemented in the C programming language and is executed from within an Arc/Info macro. Thus, Arc/Info is used to do the basic tasks like the polygon to raster conversion and the establishment of a list of population totals by region. The Arc/Info GRID file is converted into an ERDAS raster image, which is read and processed by the C program. The output from this pycnophylactic reallocation program is, in turn, converted back into Arc/Info GRID format for dissemination, mapping and analysis. ERDAS is used as an intermediate file format because the internal format is published and the files can therefore easily be read and written to by external programs. The Arc/Info export format is increasingly used as the standard format for data exchange, and we have therefore also used it to deliver the compressed files. As noted in the appendix, copyright restrictions do not permit us to the release the actual boundary coordinates.

While the preparation and interpolation of the raster data are conceptually fairly simple, several logistical problems arose as we implemented the gridding process in Arc/Info's GRID program. The two major - in the sense that they required a great deal of time - problems, and how we overcame them were:

1) Very small administrative units disappeared in the polygon to raster process, and were not assigned to any grid cell. Some of these have a large population; and

2) Gaps appeared between regional grids due to small differences in the country boundaries obtained from different data sources.

The reason for the omission of small administrative units is attributable to the discretizing effect which occurs in any polygon to raster procedure. It is generally impractical to choose a grid sufficiently fine to avoid this. But the pycnophylactic interpolation procedure was meant to be implemented with a minimum of four grid cells in each polygon. This may require an excessively fine grid, yielding too many rows and columns to be practical. Thus a modified strategy needed to be adopted. The first step in our application was to use the routines in Arc/Info's GRID module, the suite of programs that ESRI has developed to manipulate raster data. Arc/Info includes the programs required to convert from the standard vector (polygon) format to the grid (raster) format. According to the program documentation, the algorithm uses a "majority rule." That is, the program overlays a raster on the vector data, and if a part of more than one polygon cover a part of a grid cell, the value given to that cell is taken from the polygon that fills the majority of the cell. This is reasonable if only two polygons overlap a given cell, but it may not apply if several polygons overlap a cell. The documentation is in fact not entirely clear, but the process for assigning a value to a cell more likely follows a plurality rule. That is, the polygon which takes up the largest areal proportion of the cell is taken as the value for the cell. Any other polygons are ignored. All vector polygon to raster programs will have this problem. It is important that they announce when small polygons are missed in this manner.

In the present instance the areas of the five minute quadrilaterals range from eighty five square kilometers at the equator to twenty two square kilometers at the extreme latitudes of our data. In the low latitude regions with fairly high resolution administrative unit data, e.g. some of North America and especially Central America, several census units are under twenty square kilometers in area. These units do not occupy enough area in a cell for their value to be included in the gridded data. Since many of these units have high populations (i.e. the urban districts), it was necessary to derive a method for including them in the grids.

The first step was to grid the vector data at a smaller cell size. We chose 2.5 minutes, twice the final resolution, for this process because it was small enough to include most of the units omitted at five minutes but still large enough so as not to be unwieldy. For example, in North America, thirty five units were omitted at the five minute cell size and only fifteen units were omitted at the two and one-half minute cell size. It was then possible to create a separate grid containing only the missing units. In these grids, each of the units occupied only a part of one cell, but the data is of course taken to apply to the entire cell. The populations of the missing units were then merged with (added to) the original 2.5 minute grids resulting in grids that included all of the population for every administrative unit. This process is fairly simple using the relational database capabilities included in Arc/Info with the GRID merge operation. After creating the 2.5 -minute grid that included all of the units, the cell populations were aggregated to the 5 minute quadrilaterals. Since the average resolution of the original polygon units was close to eighty five square kilometers the 2.5 minute grid was considered unrealistically precise and not appropriate for distribution as the final grid. For the smoothed redistribution the pycnophylactic program performed the interpolative smoothing at the 2.5 minute resolution; the program then aggregated the resulting population grid to the 5 minute cell size. The results of this pycnophylactic interpolation formed the final smoothed product.

The process of including missing units could have been performed with 5 minute quadrilaterals instead of first gridding the vector data at the finer (2.5 minute) resolution. However, this risked the problem of overwriting cells that were unique cells for other units thereby eliminating them from the grid when incorporating the grids created from the units missing from the original grids. In other words, at the five minute cell size it is likely that several units were only large enough to occupy one cell, and it is possible that they were near or adjacent to other units that were omitted because they also were too small. When these omitted units are gridded and the grid of these units merged with the original grid, the cells of the missing units coincide with the cells representing the units small enough for only one cell. When the two grids are merged, the grid cells of the missing units could overwrite the cells for the one-cell units. This in fact occurred in Central America at the 2.5 minute resolution. Since we did not want to grid the vector data at smaller resolution because of space constraints and processing requirements, several individual solutions were required. First, the units that were omitted during the first gridding process and those omitted as a result of combining the grids described above were extracted from the Central America coverage. As expected, such units were often adjacent. The adjacent units and their populations were then aggregated and the original Central America coverage updated by replacing the individual omitted units with the larger units. This process resulted in a loss of spatial resolution for the polygons, but since the resolution of the units was anyway finer than the grid cell, it does not affect the final result.

Gaps between regions occurred because our input data of administrative unit boundaries came from a wide range of sources. There were many cases where the national boundaries did not exactly match. When we assembled the countries into regions, or received data for an entire region, we addressed these mismatches. However, when we gridded several regions at one time we had to make sure that there were no gaps or slivers in the grids due to boundary mismatches when adjacent regions were joined. This entailed adding individual cells to some regions. Such a process should be fairly simple, but in fact it required several tedious steps. First the region grids were joined to identify where the boundaries did not match and empty cells occur-red. These areas were extracted into a separate grid and the empty cells, i.e. the cells with no data, were given a temporary value and the other cells were given no data. This resulted in grids with an arbitrary identifying temporary value, (e.g. an easily recognizable number such as two), in the cells that were empty on the border between two regions. This grid was then merged with one of the region grids. This created a new grid which was the region grid with some new cells on the region boundary with the arbitrary numerical identifier (two, in the example). The GRID procedure NIBBLE was then used to give these edge cells the value of the unit closest to it using Euclidean distance. Thus, since these edge cells were often lonely individual cells, as opposed to several layers of cells, this process simply added a few cells to some of the units on the boundary between regions. Of course, when these secondary grids also omitted some units, this procedure had to be repeated. The augmented grid was then used in the population assignment and smoothing process.

SURFACES

In the previous paragraph we described some difficulties in the gridding process. Here we note some difficulties in the reallocation process. The accompanying figure shows the population densities in the raw and reallocated raster images for the whole world. In the global over-view map the difference between the smoothed and unsmoothed surfaces is barely visible because of the generalization due to the small map scale. The maps for Central America, West Africa, and Southeast Asia show the differences more clearly.

Several limitations of the source information are apparent from these images. The resolution of the administrative units used was dependent on data availability rather than the result of an ideal design. Consequently the resolution sometimes varies dramatically between as well as within countries. We tried to address this problem by varying the number of iterations in the smoothing reallocation algorithm, using the average resolution of the computational regions. The number of iterations varied from only thirty for Central America to one hundred for the former Soviet Union and the Middle East, where the units are larger. Within many regions the regional variance in resolution was still too large to completely overcome this problem. In Brazil, for example, very large units are adjacent to many small ones and no single iteration count was able to yield completely satisfactory results. A modification of the algorithm to be adaptive in the amount of smoothing might be more successful. This might require a grid size adapted to the resolution, with a corresponding small change in the computer program, followed by reaggregation to a uniform grid. A further unanticipated result appeared in the population reallocation method. When the units are relatively homogeneous in resolution the method works well and is appropriate. In cases where small units abut large ones, or the population density changes drastically from one administrative unit to the next, certain anomalies occur. For example, where urban districts are located next to rural areas, the objective of achieving maximum smoothness will "pull' nearly all of the population of the rural region tightly towards the boundary with the urban region. This is apparent in the maps, e.g., in Sumatra and central India. The problem seems exacerbated in areas where the resolution variance is greatest. Suggestions to overcome this difficulty are given in Part III of this report.

RESULTS

The final useful products from the several processing steps include:

- A piecewise continuous population surface, without any smoothing, i.e., a constant population within each administrative unit at the 5' by 5' resolution, taking into account the earth's roundness and the convergence of the meridians. Cells nearer the equator within any unit have a slightly larger population because of the earth shape. This gives a constant density to the unit.

- Smoothly reallocated population values, by 5' x 5' quadrilateral.

- Gridded population density surfaces, obtained by dividing each cell in the preceding two files by its spherical area, in square kilometers. Deflation of the true density values may occur in this process since the cells could contain ocean as well as occupied land. The accompanying table (Table III) gives an idea of the spherical cell areas. This table also shows the difference between the areas computed on the World Geodetic System ellipsoid of 1984 and the same areas computed on an equal surface area sphere of radius 6,371.007178 kilometers. It can be seen that the spherical approximation is adequate for the present purpose. The maximum differences do not exceed one half square kilometer, and amount to less than one percent.

- The administrative unit name(s) for each grid cell. This is particularly useful if one wishes to incorporate additional data by subdivisions of a country and automatically assign the values to the quadrilaterals.

- The centroid coordinates, in latitude and longitude, for all of the 19,032 polygons used. This form of the data does not allow recovery of the copyrighted administrative unit boundaries with any precision, but is useful for more generalized tasks. Included with the coordinates are the estimated population and calculated area of each polygon, and the administrative unit name(s). Population density at spot locations can therefore be calculated. This is a small data set of 19,032 records (928k bytes) and has been put onto an IBM PC diskette, along with a program to display the points on a PC screen. The display can be used to investigate the resolution of the assembled data in greater detail.

- Several maps are given in the appendix which further illustrate our results. The more important result of course is the digital data which can be used for analysis, correlation with other quadrilaterized data, for analytical modeling, or from which the user can produce alternate outputs and aggregations. It is easy, for example, to summarize and cumulate population by latitude or longitude strip.

Gridding the entire world by five minutes of latitude and longitude yields a raster of 180 x 12 x 360 x 12 which is 2160 rows by 4320 columns, or 9,331,200 cells. But less than one third of the earth is people-occupied land. Consequently we have truncated our grid at seventy two degrees north and fifty seven degrees south latitude since there are essentially no permanent residents at higher latitudes. The resulting grid is 1,548 rows by 4,320 columns of four bytes each (before compression). Approximately 30.9 percent of the 6,687,360 cells contain people. Part of Greenland exceeds seventy two degrees north and this is not included in the raster. The estimates for the eleven towns in this country are on the accompanying PC diskette, and can be placed in an enlarged raster if desired. The global rasters range in size from 35.6 Mb to 46.3 Mb before compression. Since a good proportion of the world is unoccupied, compression reduces the storage volume by over eighty nine percent.

DISTRIBUTION

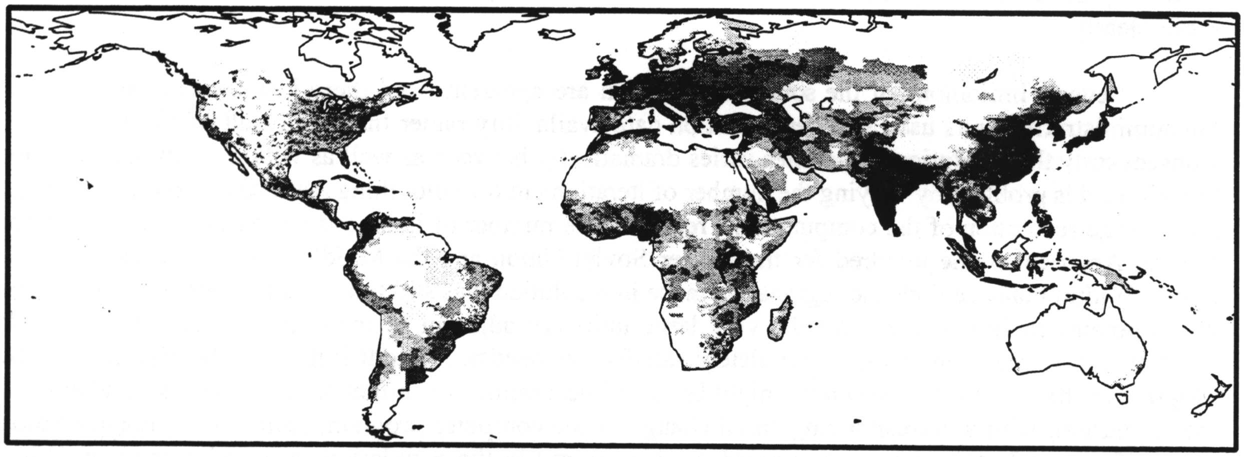
CIESIN has agreed to make the quadrilaterized population data publicly available. The current plan is to provide an ftp site, ftp.ciesin.org, placing the data in a directory called pub/data. For information contact CIESIN at one of the addresses below:

CIESIN
2250 Pierce Road
University Center, MI 48710 USA

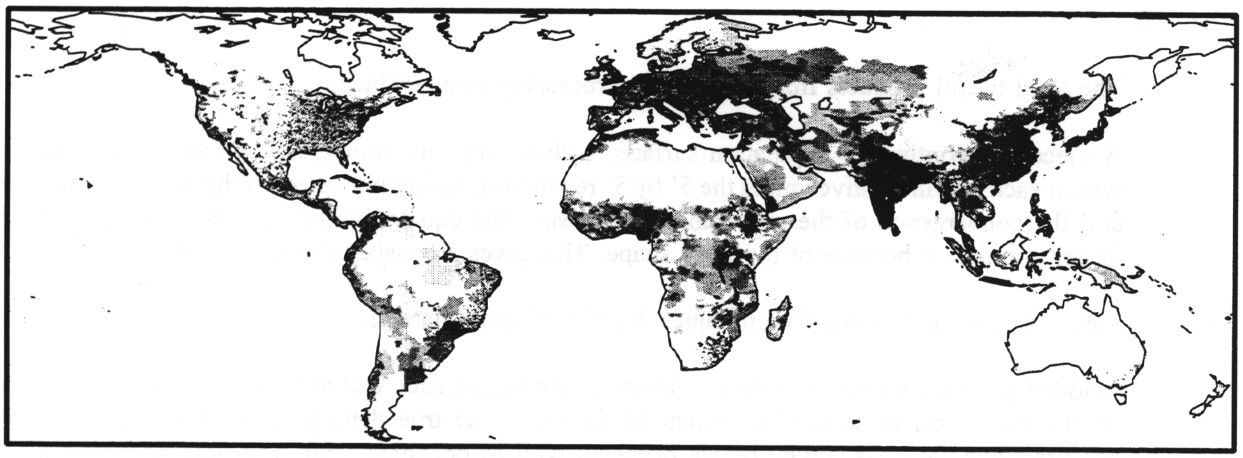
voice: (517) 797-2700
fax: (517) 797-2622
email: ciesin.info@ciesin.org

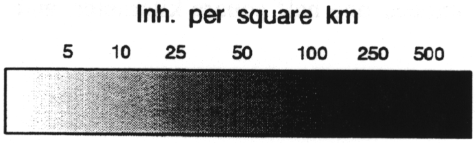
Global Population Distribution

Population Density (Unsmoothed)



Population Density (Smoothed)

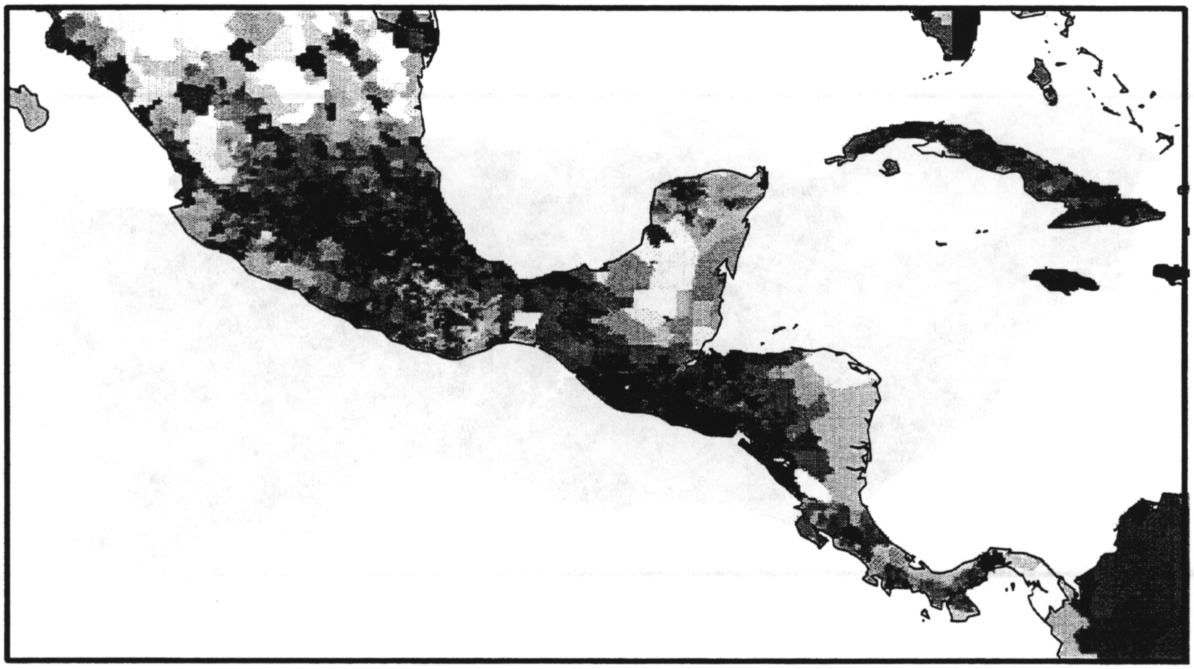




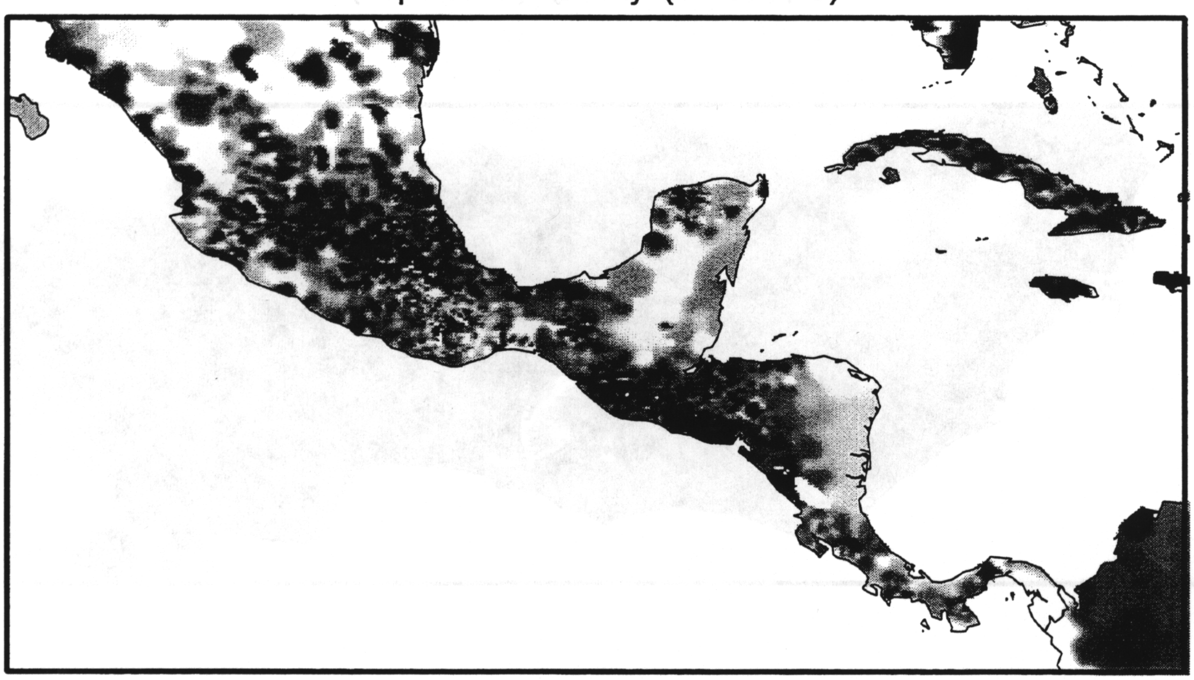
Global Demography Project
NCGIA/Geography, UC Santa Barbara

Central America

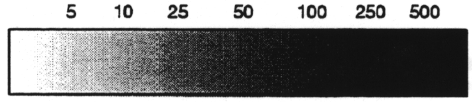
Population Density (Unsmoothed)



Population Density (Smoothed)



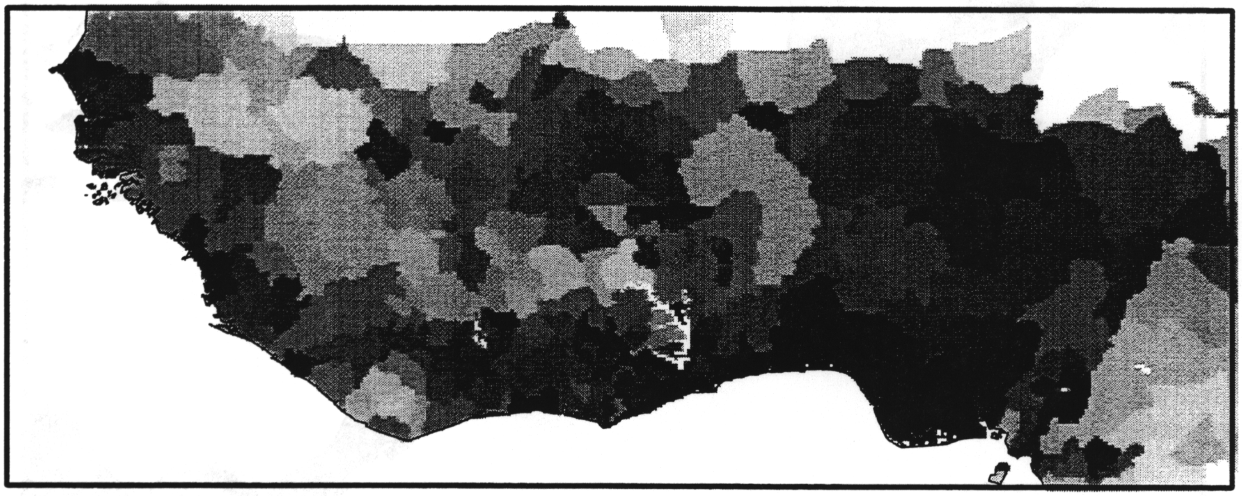
Inh. per square km



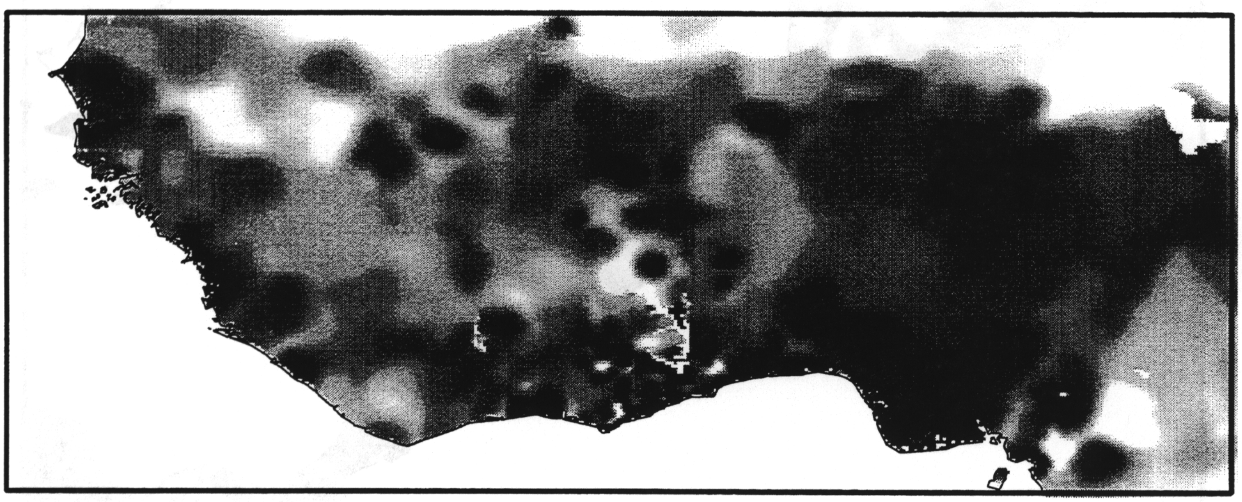
Global Demography Project
NCGIA/Geography, UC Santa Barbara

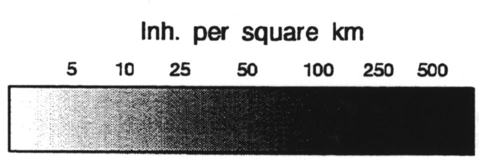
West Africa

Population Density (Unsmoothed)



Population Density (Smoothed)

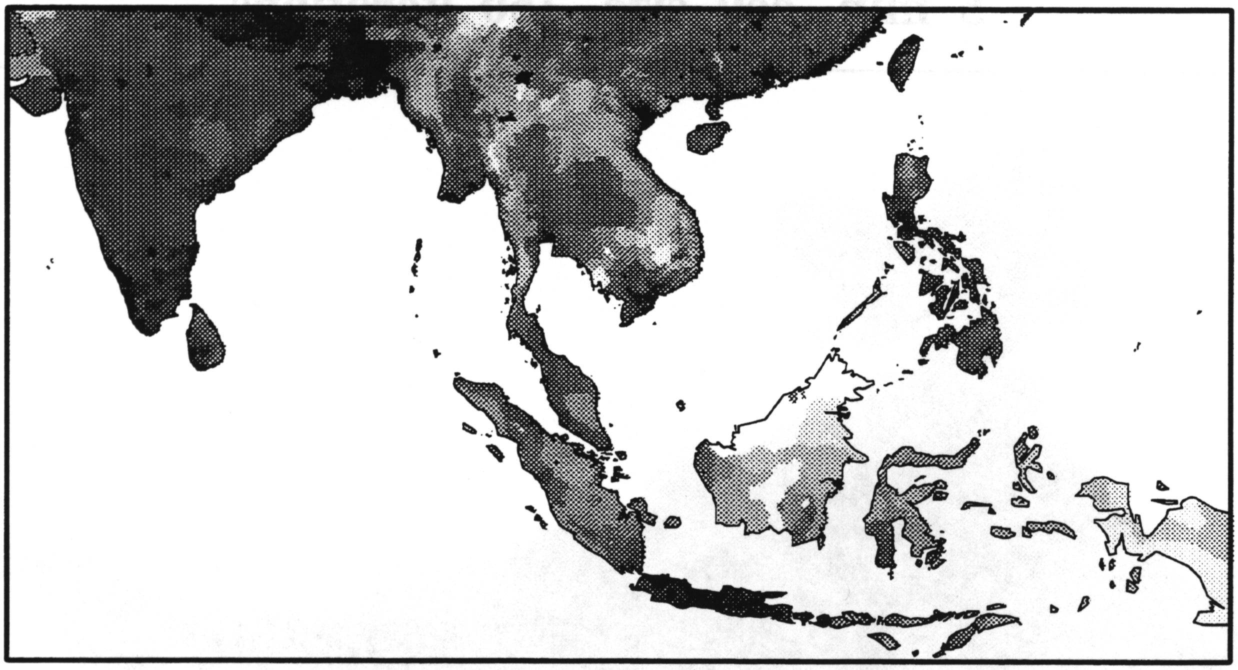




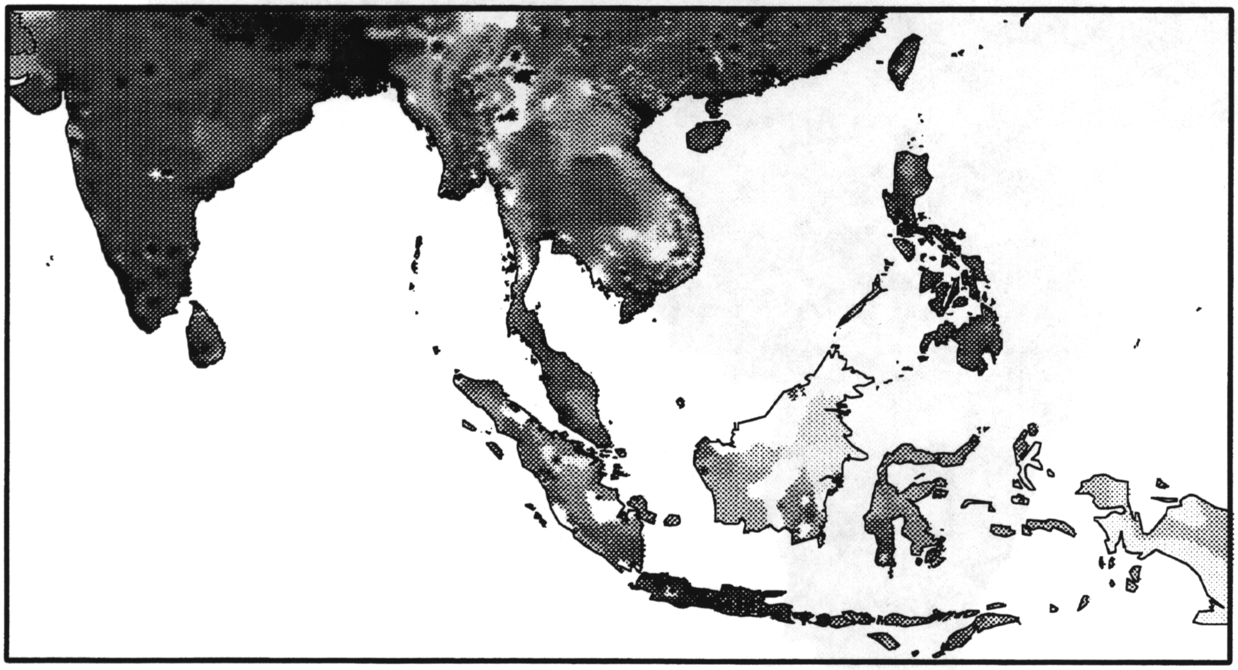
Global Demography Project
NCGIA/Geography, UC Santa Barbara

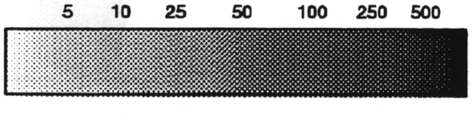
Southeast Asia

Population Density (Unsmoothed)



Population Density (Smoothed)





Global Demography Project
NCGIA/Geography, UC Santa Barbara

Pycnophylactic Interpolation
5 min. cell size, 150 iterations

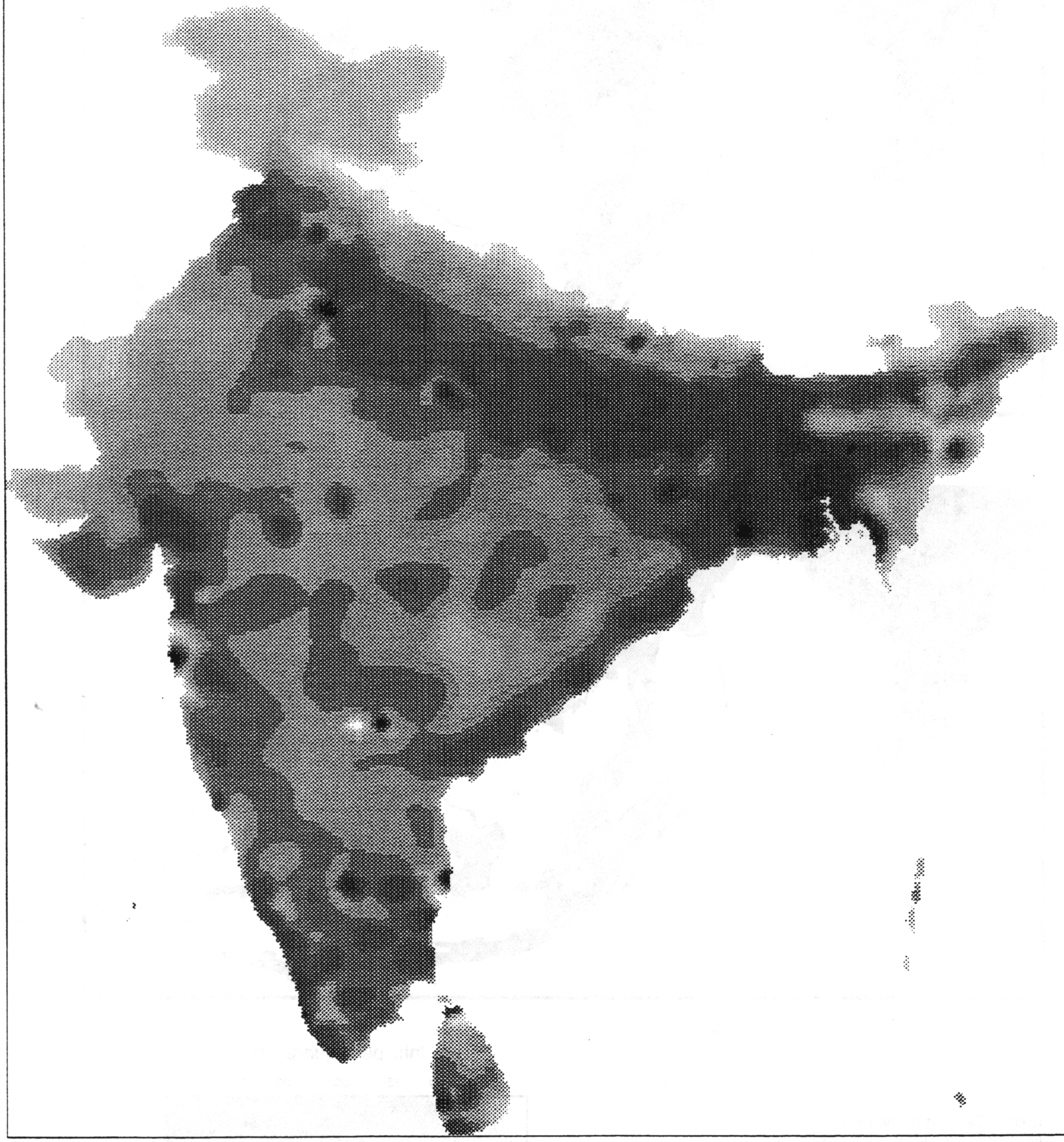


TABLE III

Difference in quadrilateral area using the exact ellipsoidal formula and using a mean radius sphere, in square kilometers. Based on a quadrangle area of 5 minutes in latitude and 5 minutes in longitude.

Latitude (Degrees)	Ellipsoid Area WGS 83	Spherical Area	Diff. sq km qerence sq km square kilometer s kmkm	Percent Diff.	E-W km Distance Distance
0.00	85.480	85.863	-.384	-.449	9.266
0.00					
5.00	85.158	85.531	-.374	-.439	9.230
10.00	84.204	84.548	-.344	-.408	9.124
15.00	82.625	82.922	-.296	-.358	8.949
20.00	80.430	80.664	-.234	-.291	8.705
25.00	77.631	77.792	-.162	-.208	8.395
30.00	74.245	74.329	-.083	-.112	8.021
35.00	70.295	70.299	-.004	-.006	7.587
40.00	65.805	65.735	0.070	0.107	7.094
45.00	60.806	60.670	0.136	0.223	6.547
50.00	55.332	55.114	0.188	0.340	5.951
55.00	49.422	49.198	0.224	0.452	5.309
60.00	43.118	42.878	0.241	0.558	4.627
65.00	36.469	36.231	0.238	0.653	3.910
70.00	29.526	29.308	0.127	0.736	3.163
75.00	22.342	22.163	0.179	0.802	2.392
80.00	14.976	14.849	0.128	0.851	1.602
85.00	7.487	7.421	0.066	0.881	0.801

Latitude refers to the southwest corner of the 5' by 5' quadrangle. The radius of 6371.007178 kilometers gives a spherical surface area of 510,065,621 km^2 equal to that of the WGS 1984 ellipsoid.
Authalic latitude has not been used.

PART III: THE NEXT STEP

The global population information described in the previous portion of this report can be considered an exploratory reconnaissance designed to stimulate interest worldwide. Now that this preliminary but basic database has been established it can be widely circulated and its usefulness determined. And we have demonstrated the feasibility of this type of work. It will hopefully lead to follow-up projects aimed at improving and expanding the current effort. Improvements can be made by incorporating more up-to-date information as it becomes available, by filling in areas for which we have not been able to obtain sufficient detail. Higher resolution data are desirable for several parts of the world, particularly the former Soviet Union where we have obtained only oblast data, and several countries in South America. Other desirata have been mentioned throughout this report, including getting more demographic components. A critical need is for a maintenance organization. One model would be to have local authorities or researchers add data to the 5' by 5' cells and to keep them up to date, using a world wide system of social observatories.

We describe here several extensions which we have considered. These include error estimation, the incorporation of urban populations, better reallocation procedures using additional auxiliary information, extension to include movement modeling, and data compaction schemes. We do not comment on the need for correlation with other items of global concern such as environmental degradation, etc. This need is rather obvious.

DATA SET COMPARISONS

Given the difficulties of obtaining population data at subnational levels in many countries of the world, we have made no attempt to estimate the magnitude of the likely errors. Such work is beyond the current resources available to us. But, during the course of this study we have discovered a few alternate collections of demographic data which might help alleviate some of these concerns. We believe that comparison of these with those we have obtained might shed light on the reliability of our methodology. In particular, the Center for International Research in the US Bureau of the Census has recently made available their estimates of population by 10' by 20' cells for some twenty countries. Our data can easily be aggregated to this level. Of course the problem of synchronicity remains. In addition, the Goddard Institute for Space Systems (Matthews 1983) has assembled world population by one degree quadrilaterals; Some enhancements to this tabulation have been performed by the Oak Ridge National Laboratory, and these are available, but nearly a decade old. Several countries, including the United States, have population data at a much finer resolution than we have attempted for the world. Some countries have demographic data by dwelling unit, recorded in coordinates to the meter or decameter level! These figures could be aggregated to the medium resolution we have assembled and used as an estimate of the errors introduced by our redistribution. A few countries even already have such aggregations and these can be compared directly with our assemblage in future work. A useful investigation would be to examine ways in which data assembled at such micro levels have been used, and whether this is warranted for larger areas of the world.

URBAN POPULATIONS

In some cases it may be possible to distinguish between urban and rural populations, treating them separately. When an urban population count is available by city and if the latitude & longitude of the center of the cities are known then it is possible to make an estimate of the population at various distances from the center of the city. Thus the people can be assigned to the cells of a lattice around the city, after which the rural population portion can be allocated separately. This should provide a better estimate of the population arrangement than simply using the city centroid as a point of large population. Of course it would be even better to have the urban area digitized as an individual polygon; this was in fact often the case in our second level data. The bad news is that we have not been able to discover any source of comprehensive worldwide city demographic information with latitude and coordinates for our use in those cases where this would have been helpful. We have found one file, apparently based on United Nations data, predating 1992, of 2763 cities with populations and latitude, longitude coordinates. The population in these cities sums to 1,334,783,002 people. We have not used this data in our compilation because much of the same information is captured in the administrative units. A more comprehensive file, if it existed, could however be used to compare with our assemblage, or even to enhance it. The good news is that there is empirical evidence telling us how this might be done. It seems that relating the population of a city to the geographic area which it covers can be done with some degree of confidence (Tobler 1969b). Although the correlations are high this must still be used with caution since the functional city and the political city often do not coincide. The land area estimate pertains to the functional city, but population counts are most often based on political cities.

The two empirical regularities on which the following discussion is based pertain to the city sizes, as a function of population, and to the arrangement of people within a city. The idealized circular city size is based on the observation that the land area of a city can be predicted with considerable accuracy by knowing the number of its residents. Mathematically stated, City Area = $a(\text{Population})^b$, or taking logarithms, Ln Area = Ln a + b Ln P. From this an assumed circular city size of equivalent area can be computed. The total population density of a city, in persons/kM2, can also be estimated from this empirical formula for city area. We also find that the change in area, or radius assuming a circular city, due to an increase of one person is $dA = \alpha\, dp^{\beta}$ or $dR = \theta\, dP^{\gamma}$, where the new coefficients α, β, θ, and γ can easily be determined from the above formula. It is then seen that the average density increases with city size, but that the influence of an individual on the total area or radius of a city decreases with increasing city size. Perhaps the individuals' political influence also varies in this manner.

For the United States, using P for the total population and R to denote a city radius in kilometers, the empirical relation is $R = 0.035\ P^{0.44}$. In some other cultures the scaling coefficient, a, is generally smaller, resulting in more compact towns, but the exponent, b, appears stable. The empirical observations from which this relation was inferred cover settlement sizes ranging from 150 people to over a million persons with very high correlations. A similar rule is used, in inverted form, by some archaeologists to estimate the former population of excavation sites. The constant of proportionality, a, may also change with time. In the United states it appears to have changed in 1945 when the impact of the automobile affected suburbanization (Tobler 1975).

The next step is to choose a model of the arrangement of the population around the center of the city, assumed circular. Once this is done the population can be reallocated to the geographic area of the city. The density at the center of the urban area can also be estimated from the model by setting the distance from the center to zero. Since circular symmetry holds we use polar coordinates. Using r for the distance from city center and D(r) for the population density out from the city center, the following models are obtained from D = f(r) by setting

$$P = \int_0^{2\pi}\int_0^{R} D\,dr\,d\theta = 2\pi\int_0^{R} Dr\,dr \quad , \quad \text{and } D = 0 \text{ (or } D = 1) \text{ at } r = R.$$

In words, the total population of the city must be contained within its radius (Tobler 1969a). When the population density remains non-zero out to an infinite radius (as in the exponential and Gaussian city models, below) the city area is arbitrarily truncated at a population density of one person per square kilometer. The models are listed in increasing order of complexity, and probably realism:

Cylindrical city	$D = P/\pi R^2\ ,\ r < R$	(constant density)
Conical city	$D = 3P/\pi R^2 - 3Pr/\pi R^3$,	$r < R$ (linear decline)
Parabolic city	$D = 6P(R - r)^2/\pi R^4$	$r < R\ ,\ \partial D/\partial r = 0$ at R
Cosine city	$D = \pi P(1 + \cos[\pi r/R])/\{R^2(\pi^2 - 4)\}$,	$r < R\ ,\ \partial D/\partial r = 0$ at R
Exponential city	$D = Ae^{-br}$	$(D_R = 1)$
Gaussian city	$D = Ae^{-br^2}$	$(D_R = 1)$

There is disagreement in the literature as to which model is most appropriate, and, as models, none can be expected to be perfect. Nor are cities really circular, although the equations predict the land area and total density quite well. Other models, including gamma functions and other more complicated types yielding lower populations in the city center exist (e.g., Vaughn 1987); some versions result in different urban densities in different azimuths (Medvedkov 1965; Haynes 1973) and for variations over time (Bussiere 1972; Newling 1973). To get the total population numbers it is necessary to convert the densities to a count, inverting the foregoing equations. Using these models one can assign the urban population to a restricted (circular) subset of the area and spread the rural population over the remainder. Adjustments may need to be made at political boundaries, and overlaying a coastline file would be useful in avoiding the assignment of people to water areas and might require a corresponding adjustment to obtain the total land area suggested by the model. In these cases the theoretical radius also might suggest areas for potential future city growth by infilling.

This method of reallocating urban population, using a kernel function to convolve the data, is quite analogous to the one used in statistics to assign two dimensional densities to scattered observations (e.g., Silverman 1986; Scott 1992). Examples used to distribute urban populations are available in the work of Honeycutt & Wojcik (1990) for the United States, and by Bracken & Martin (1989) and Martin & Bracken (1991) for a region in Great Britain. The kernel is generally chosen to mimic the (theoretical) distribution of population within a city using one of the foregoing models. Often the urban centroid coordinates are not available, though they may be estimated with sufficient accuracy by examining maps, or by choosing a dominant city within each region. Complete inventories of world city populations by latitude and longitude have not been located. This method has therefore not been used though it would seem to offer some promise.

SMART POPULATION REALLOCATION

The smoothing redistribution used to create the population surface maintains the correct population totals for each polygon and includes the geographic insight that neighboring places have similar values. But it is rather mechanical, and, as previously noted, also has a curious consequence. When a high density population unit is adjacent to one of sparse population, the smoothing pulls virtually all of the population of the sparsely populated region to the border adjacent to the densely populated unit, leaving the remainder

of the sparsely populated unit completely devoid of population. Such a condition may be realistic, but looks strange on density maps - see, for example, the region around Hyderabad in central India (page 24). Thus the method used for the spatial disaggregation of the population needs to be improved. One approach is to obtain further guidance from geographic theory; for example, the cited interaction potential of population and the literature on urban density gradients. Thus we believe that the reallocation can be greatly improved by using additional data on factors that influence the population arrangement. These factors would include the location and size of towns and cities, major infrastructure (roads, railroads), which of course is itself influenced by the population, and natural features (rivers, uninhabitable areas, protected and wilderness areas, etc). The feasibility of this approach has been shown in a study by UNEP/GRID in which a population density surface for the African continent was produced and subsequently published in UNEP's Global Atlas of Desertification; see Deichmann and Eklundh (1991) for a description of the methodology. In that project, efficient use was made of the capabilities of GIS to integrate heterogeneous data within a consistent framework. This approach was subsequently adopted and modified for producing population maps of the Baltic States at one square kilometer resolution (Sweitzer and Langaas 1994) and for the European continent at 10' resolution (Veldhuizen et al 1995). To apply this to the entire world requires overlays of the infrastructure and physical geography and these have not been used in the present project. But this work could now be facilitated by the availability of a consistent global base map, namely the Digital Chart of the World (DCW) or the digital world topography (at 5' by 5'), both now available on compact disks (CDs). Taking into account secondary information is referred to in some interpolation literature as co-Kriging, and is an obvious approach as long as the additional information is available and its relation to the primary variable known. Thus earth satellite sensors capable of recording visible emissions at night, used judiciously, also offer promise as a means of detecting, correcting, or calibrating global population concentrations.

MOVEMENT MODELING

The immigration-emigration component of demography has not been included in the foregoing discussion. To remedy this a possible model for migration is given here. It assumes that the information is given by latitude/longitude quadrilaterals. As an abstract model it should work for refugees, and for global trade too. Since all change requires movement (of information, of people, of credit, or of material) more stress should be placed on the global measurement of movement. Most international agency data collections are of state variables rather than of flow variables, to use the language of systems analysis. Tables of industrial or agricultural commodity trade, when available, can be treated in a manner similar to that described below for migration tables, as can more general communication information. The geographic movement of ideas and global monetary flow are as important as wind and weather movements. Getting the data, promptly and continuously, is the hard part. Aside from this the details are not difficult.

Movement phenomena can be recorded in several ways, including traffic counts on links, or by noting boundary crossings. The latter result in "From-To" tables. In each instance a time interval, usually fixed, is also specified. A short list of such interaction tables would include

Face to face personal contacts
Business contacts
Commodity trade
Telephone calls
Migration
Letters

In the first example a global table including everybody in the world would be circa 5.6×10^9 by 5.6×10^9 in size, but extremely sparse, with a complexity and connectivity approaching that of the human brain. In the other cases we assume that geographic locations identify the rows and columns of the table, and the entries in the body of the table consist of a count or measurement of the interaction between the places. When there are N places these tables are N 2 in size. Each usually records the events during some fixed time interval. For trade purposes countries typically record only one column (incoming items), occasionally one column and one row (both incoming and outgoing), of such a table with the data applying to their interests, though they may have these for many items. International cooperation is required to obtain complete from - to estimates. A world table of migration or trade using 200 countries needs to contain 40,000 numbers, though many (most?) might be zero. The typical table is asymmetric, announcing uneven movement in the two directions. There are several categories of spatial model used for the analysis of such tables. Only one such possibility is described here, for migration.

People move for many reasons, some obvious, some idiosyncratic. The model abstracts these and asserts that migration results from some dissatisfaction, discouragement, repulsion, rejection, or push from one's current location and an offer, opportunity, enticement, allure, temptation, fascination, or pull from some other region, modulated by the difficulty of transfering from one place to the other (Dorigo and Tobler 1983). Algebraically this can be written as

$$M_{i,j} = (R_i + E_j) / d_{i,j} \;,$$

where the movement (M) takes place between regions, N in number, indexed by i and j. The pushes (R) and pulls (E) are real numbers which can take on positive or negative numerical values and are to be estimated. The cost of movement from place i to place j is contained in the dij term, for which great circle or road distance is often an adequate surrogate.

Given an empirical N by N from-to movement table and observed distance disutilities the numerical values of the push and pull factors can be calculated for each place. Contrast this with the usual procedure of doing it the other way around, using an *a priori* postulation of causation with a regression model to assess parameters. In the model being discussed it in fact turns out that these pushes and pulls can be estimated from the marginals (sum over rows and over columns) of the table alone. This is obviously a big savings in data collection effort; it requires only two numbers for each place, for a total of 2N values not the N^2 numbers needed for the entire from-to table. The pushes and pulls can be shown to be the Lagrangians of a constrained quadratic optimization, an interpretation which will make more sense to economists. The computed potentials (the R & E values) are then shadow prices. Since their numerical value depends on interaction between all of the places they have considerable spatial and temporal inertia, and seem candidates for forecasting models.

Interestingly the attractivity (the algebraic difference between the pull and push) of a place can be computed directly if one knows only the difference of the marginals, i.e., the net change at a place due to the movement. This requires even less data, only N numbers, but these may now be positive or negative (Tobler 1981).

Now suppose that the number of places increases. The migration table becomes larger and larger. What happens in the limit, as N approaches infinity? This suggests looking at the spatially continuous case, where every location, now labeled by xy coordinates, has a spatially continuous outmigration Ox,y and inmigration Ix,y, and a push Rx,y and pull Ex,y. The difference between the in and out movement is the net change, Cx,y = Ix,y – Ox,y. Assumptions equivalent to those given above now lead to Poisson's equation

$$\frac{\partial^2 A_{x,y}}{\partial X^2} + \frac{\partial^2 A_{x,y}}{\partial Y^2} = C_{x,y} \;.$$

The potential function Ax,y (= Ex,y – Rx,y) computed as the solution to this partial differential equation is of course the attractivity. It's the same attractivity as mentioned above, but now it is a continuous function Ax,y of position. It can be represented by contours on a map. Once Ax,y has been found then a continuous flow (vector) field can be defined by

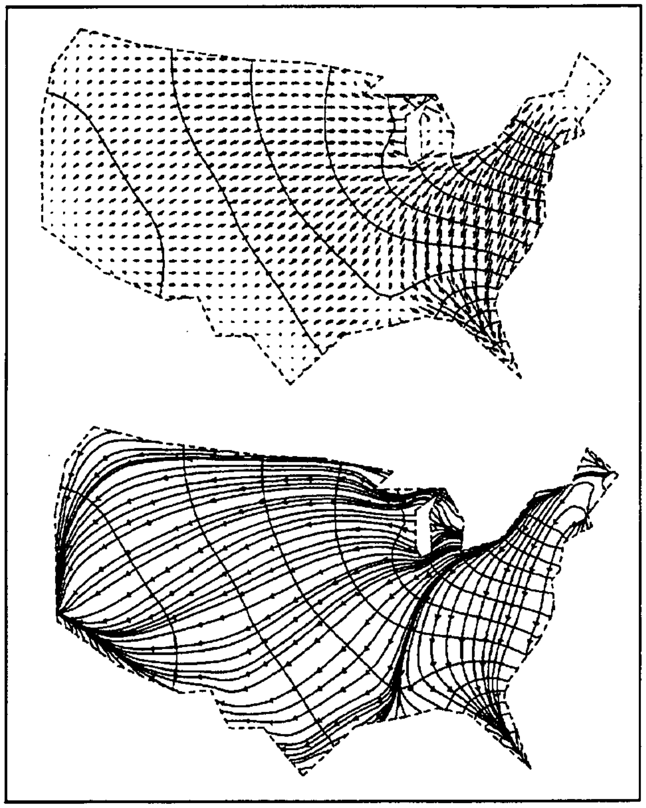
$$\vec{v}_{x,y} = \nabla A_{x,y} \;,$$

the gradient field to the contour map. A streakline map is perhaps more dramatic, and shows the partitioning into flow regions quite nicely. The actual computation for Ax,y uses a finite difference approximation. The figure shows the model applied to the 1965-1970 net migration by state for the contiguous United States. In the upper map are the contours of the attractivity, with the gradient vectors superimposed. The lower map converts the gradient vectors to streamlines shown on top of the attractivity potential.

If simultaneous two-way movements are considered then the previous Poisson equation must be generalized. It turns out that a coupled pair of Helmholz equations will do the job. These are

$$\nabla^2 E_{x,y} + 4T_{x,y} - I_{x,y} = 0$$
$$\nabla^2 R_{x,y} + 4T_{x,y} - O_{x,y} = 0$$

Subtracting the second equation from the first gives back the Poisson equation for the attractivity. Adding the two equations gives a single Helmholz equation for the turnover potential Tx,y for the two-way movement.



The Poisson equation is the solution to a problem which minimizes the integral

$$\iint\{\frac{\partial A_{x,y}^2}{\partial X} + \frac{\partial A_{x,y}^2}{\partial Y} + \alpha^2 A_{x,y}^2 + 2\beta A_{x,y}\}\ dXdY$$

and this is equivalent to an assumption that the movement is in some sense efficient - parallel to the constrained quadratic optimization - and takes place only between adjacent places. Thus movement from California to Maine must, along the way, pass thru New York state. The necessary boundary condition for the partial differential equations is that no moves can leave the system, which the same as saying that the ins and the outs of the movement table must sum to the identical value so that everything is accounted for. Real tables often do not satisfy this criterion due to inevitable problems in working with empirical data. Some adjustment procedure is then applied to satisfy this constraint. And a clever shipper from California might travel through Canada on the way to Maine, violating one form of the boundary condition of the model, or use airfreight and thus bypassing the two dimensional assumption of the model.

All of this has been formulated for a plane region. On a sphere Poisson's equation becomes slightly more complicated and the necessary inputs become functions of latitude and longitude. The appendix gives the changes needed to apply this model to a sphere, allowing calculation of the model parameters. Estimating the model numerically is easiest if the data are given at regular spatial intervals. Programs exist to estimate and draw contours or vector fields on pictures of a sphere, and mathematical functions can also be fit to the results, our next topic.

DATA COMPACTION

Verified theory is of course the best data compaction method, but for numerical data the classical method, since the time of Gauss, used by geophysicists for describing continuous variables distributed over the surface of the earth is that of spherical surface harmonics, the spherical equivalent of Fourier series. This has also been reported for the arrangement of population on the earth (Tobler 1992). Non-scalar demographic change components, such as migration, may be represented by vector spherical harmonic fields, and similar techniques may be used for trade data (Raskin 1994). But it is not apparent that the arrangement of people completely satisfies the harmonic assumption. Consequently, less restrictive spherical wavelets have been investigated (Tobler 1992, unpublished). Spherical splines have also been used to fit spherical distributions, including population (Wahba 1981; Dierckx 1984; Slater 1985). This remains an area for further development and research applied to human relations.

BIBLIOGRAPHY of references cited:

Bracken, I., & D. Martin, 1989, "The generation of spatial population distributions from census centroid data", *Environment & Planning*, A: 21:537-543

Bussiere, R., Ed., 1972, *Modeles Mathematiques de Repartion des Populations Urbaine*s, Paris, Centre de Recherece d'Urbanisme.

Chengrui, L., et al, eds., 1987, *The Population Atlas of China*, Hong Kong, Oxford University Press.

Clarke, J., and D. Rhind, 1992, *Population data and global environmental change*, International Social Science Council, Programme on Human Dimensions of Global Environmental Change, Paris, UNESCO.

Clark, C., 1951, "Urban Population Densities", *J. Royal Stat. So*c., A, 114:490-496.

Danko, D., 1990, "The Digital Chart of the World Project", *GIS/LIS Proceedings*, Anaheim; 392-401.

Deichmann, U., and L. Eklundh, 1991, "Global digital datasets for land degradation studies: a GIS approach," *GRID Case Studies Series No. 4*, Global Resource Information Database, Nairobi, United Nations Environment Programme.

Dierckx, P., 1984, "Algorithms for Smoothing Data on the Sphere with Tensor Product Splines", *Computing*, 32:319-342.

Dorigo, G. & W. Tobler, 1983, "Push-Pull Migration Laws", *Annal*s, Assoc. Am. Geographers, 73(1):1-17.

Dureau, F., O. Barbary, A. Michel, & B. Lortic, 1989, *Area Sampling from Satellite Images for Socio-Demographic Surveys in Urban Environments*, Paris, Institut Francais de Reserche Scientific pour Developpement et Cooperation.

Fowler, C., and G. Barnes, 1992, "Modeling the relationship between deforestation and malnutrition in Zaire using a geographic information system (GIS) approach," *ASPRS/ACSM/RT 92 Technical Papers*, Vol. 3, GIS and Cartography, Washington, D.C., American Society for Photogrammetry and Remote Sensing; 24-33.

Freedman, P., 1993, "Counting Countries: What makes countries count?". *The Geographical Magazine*, March; 10-13.

Goodchild, M., L. Anselin, & U. Deichmann, 1993, "A framework for the areal interpolation of socioeconomic data", *Environment & Planning*, A, 25:383-397.

Haaland, C., & J. Heath, 1974, "Mapping of Population density", *Demography*, 11:321-336.

Hagerstrand, T., 1955, "Statistika Primaruppgifter, flygkartering och 'dataprocessing'-maskiner, Ett kombineringsprojeckt", *Svensk Geografisk Arbok*, Lund, 1955.

Haynes, K., & M. Rube, 1973, "Directional Bias in Urban Population Density", *Annals*, Assn. Am. Geographers, 63(1):40-47.

Hickman, B.., 1983, *Global international economic models: selected papers from an IIASA conference*, Amsterdam, North Holland.

Honeycutt, D., and J. Wojcik, 1990, "Development of a population density surface for the conterminous United States," *Proceedings GIS/LIS*, Vol 1, Anaheim; 484-496.

Keyfitz, N.,and W. Flieger, 1990, *World Population Growth and Aging*, Chicago, University Press.

Lee, Y, 1989, "An allometric analysis of the US urban system 1960-1980", *Environment and Planning*, A, 21:463-476.

Liu, Y., 1987, "National Population Atlas of China", *13th International Cartographic Conference Proceedings*, Morelia, Vol III:481-492.

Martin, D., and I. Bracken, 1991, "Techniques for modelling population-related raster databases", *Environment and Planning*, A, 23:1069-1075.

Martin, G., 1992, *Final report of the CGIAR/NORAGRIC/UNEP meeting on digital data requirements for GIS activities in the CGIA*R, Arendal, Norway, September 21-24, 1992.

Matthews, E., 1983, "Global vegetation and land use: new high-resolution data bases for climate studies", *J. Climate Appl. Meteoro*l., 22:474-487.

Maury, M., 1851, *Whale Chart*, National Observatory, Washington D.C.

Medvedkov, Y.,1965, "Ekonomgeograficheskaya Uzychennosty Rainov Kapitalisticheskogo Mira", *Prilojeniya Matematiki v Ekonomucheskoi Geographii*, 2, Moscow, Institute for Scientific Information, pp 110-120.

Muehrcke, P., 1966, *Population Slope Maps*, MA Thesis, Ann Arbor, University of Michigan, Department of Geography.

National Research Council, 1994, *Science Priorities for the Human Dimensions of Global Change*, Washington D.C., National Academy Press.

Newling, B., 1973, "Urban Growth & Spatial Structure: Mathematical Models and Empirical Evidence", *Ekistics*, 36:291-297.

Park, A., 1990, "Smooth Pycnophylactic Areal Interpolation and the Estimation of Missing Data", MS Thesis, Leicester, Leicester University, Department of Geography.

Raskin, R., 1994, "Spatial Analysis on the Sphere: A Review", Technical Report 94-7, Santa Barbara, NCGIA, 44 pp.

Rhind, D., I. Evans, & M. Visvalingam, 1980, "Making a National Atlas of Population by Computer", *The Cartographic Journal*, 17:1-11.

Rogers, A., 1985, *Regional Population Projection Models*, Beverly Hills, Sage.

Rylander, G., 1986, "Areal Data Reaggregation", MA thesis, East Lansing, Michigan State University, Department of Geography.

Scott, D., 1992, *Multivariate Density Estimation*, New York, J. Wiley.

Silverman, B., 1986, *Density Estimation for Statistics and Data Analysis*, New York, Chapman & Hall.

Slater, P., 1993, "World Population Distribution", *Applied Mathematics and Computation*, 53(2-3):207-223.

Stern, P., O. Young, D. Druckman, (Eds), 1992, *Global Environmental Change: Understanding the Human Dimension*, Washington DC., National Academy Press.

Sweitzer, J., S. Langaas, 1994, *Modelling population density in the Baltic States using the Digital Chart of the World and other small scale data sets*, technical paper, Stockholm, Beijer International Institute of Ecological Economics.

Taylor, J., 1989, "The World Maps Set", Indiana University of Pennsylvania, Department of Geography and Regional Planning.

The World Almanac and Book of Facts 1995, Mahwah, Funk & Wagnalls, 976 pp.

Thomas, R., 1990, *Spatial Epidemiology*, London, Pion.

Thomas, R., 1992, *Geomedical Systems: Intervention and Control*, London, Routledge.

Tobler, W., 1969a, "The Spectrum of U.S. 40", *Papers*, Regional Science Assn., XXIII:45-52.

Tobler, W., 1969b, "Satellite Confirmation of Settlement Size Coefficients", *Area*, I:30-34.

Tobler, W., 1970, "A computer movie simulating urban growth in the Detroit region", *Economic Geography*, 46(2):234-240.

Tobler, W., 1973, "Regional Analysis: Time Series Extended to Two Dimensions"; *Geographica Polonica*, 25:103-106.

Tobler, W., 1975, "City Sizes, Morphology, and Interaction", Internal Report WP-75-18, Laxenburg, IIASA.

Tobler, W., 1979a, "Smooth pycnophylactic interpolation for geographical regions", *J. Am. Stat. Assn*., 74(367):519-530.

Tobler, W., 1979b, "Cellular Geography"' pp 379-389 of S. Gale & G. Olsson, eds., *Philosophy in Geography*, Dortrecht, Riedel.

Tobler, W., 1981, "A Model of Geographic Movement," *Geographical Analysis*, 13(1):1-20.

Tobler, W., & Z. Chen, 1986, "A Quadtree for Global Information Storage", *Geographical Analysis*, 18(4):360-371.

Tobler, W., 1988, "Resolution, Resampling, and All That", in H. Mounsey & R. Tomlinson, eds., *Building Data Bases for Global Science*, London, Taylor & Francis; 129-137.

Tobler, W., 1990a, "GIS Transformations, *Proceedings*, National Computer Graphics Assn, Vol I., Anaheim; 163-166.

Tobler, W., 1990b, "Frame Independent Spatial Analysis", in M. Goodchild, ed., *Accuracy of Spatial Data Bases*, London, Taylor and Francis; 115-122.

Tobler, W., 1992, "Preliminary representation of world population by spherical harmonics", *Proceedings of the National Academy of Sciences of the United States of America*, 89(14):6262-6264.

United Nations Development Programme, 1990, *Human Development Report 1990*, New York, Oxford University Press.

United Nations Population Fund, 1992, *The State of World Population*, New York, United Nations

United Nations Population Division, 1993, *World Population Prospects*, 1992 Revision, New York, United Nations.

Vaughn, R., 1987, *Urban Spatial Traffic Patterns*, London, Pion.

Veldhuizen, J. van, R. van de Velde, & J. van Woerden, 1995, "Population mapping to support environmental monitoring: some experiences at a European scale," *Proceedings*, Sixth European Conference on Geographical Information Systems, Utrecht, EGIS Foundation.

Wahba, G., 1981, "Spline Interpolation and Smoothing on the Sphere", *SIAM J. Sci. Stat. Comp*., 2:5-16.

Warntz, W., 1965, *Macrogeography and Income Fronts*, Philadelphia, Regional Science Research Institute; 89-113.

Willett, B., ed., 1988, *The New Geographical Digest*, London, George Philip.

Witthauer, K., 1969, *Verteilung und Dynamik der Erdbevoelkerung*, Leipzig, H. Haack.

World Resources Institute, 1990, *World Resources 1990-1991*, New York, Oxford University Press.

K. Zachariah, and M. Vu, 1988, *World Population Projections*, Baltimore, World Bank & Johns Hopkins University Press.

APPENDICES

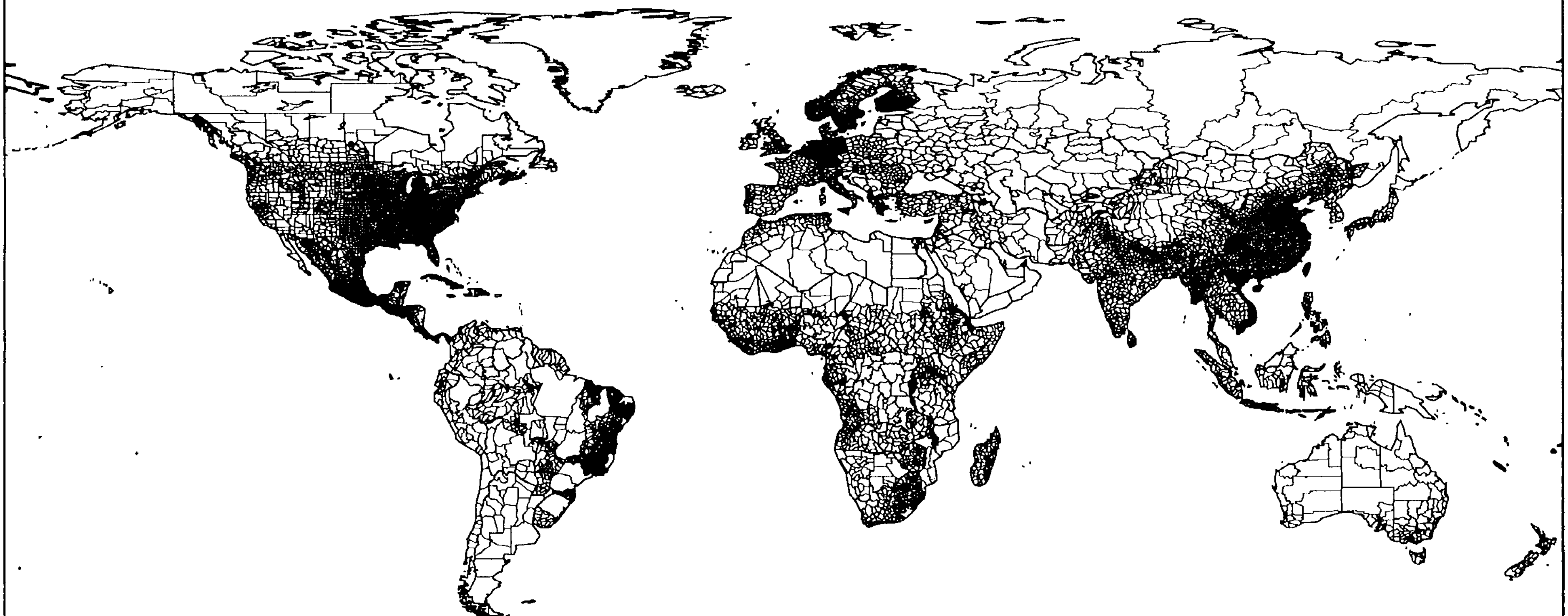
Maps of the several regions, followed by tables for each region. Then sources of population data, followed by sources of boundary coordinates. The maps are on the plate caree projection, which is not equal area. The regional subdivision is:

Americas
Europe
Africa
Asia
Oceania

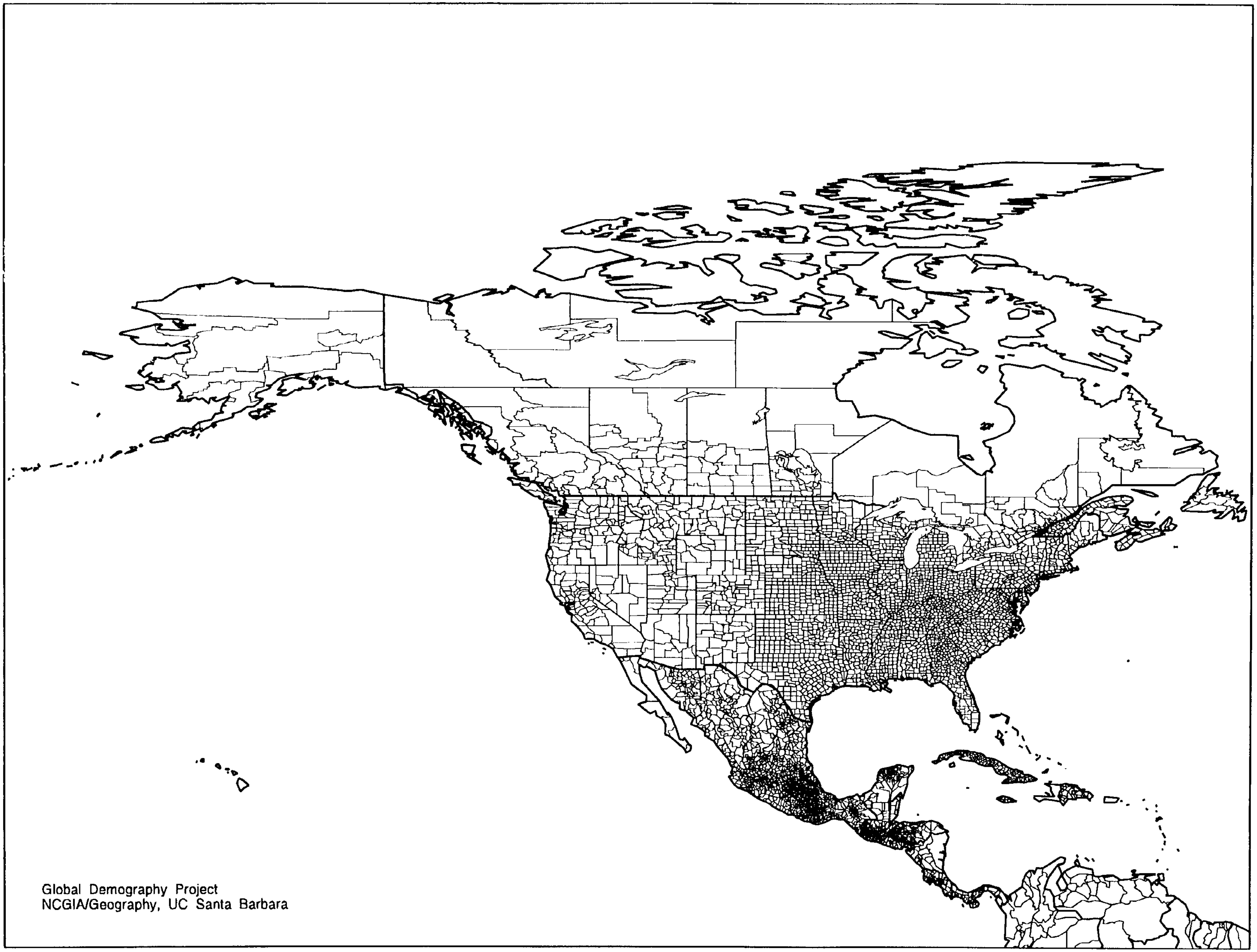
Smooth mass preserving areal data redistribution on a sphere.

IBM compatible PC diskette with the following files:

README	Table of contents.
CENTROID.LST	19,032 centroids; Lon, Lat, Pop, square km., Names
COUNTRYS.ABB	Country abbreviations.
CENSUS.LST	Number of units & resolution, by country.
REGIONS.CEN	Centers of 2,463 regions; Lon, Lat, Pop, Area, Names(s).
COUNTRY.CEN	Centers of 217 countries; Lon, Lat, Pop, Area, Name.
HIGHDENS.LST	Location of 59 areas with population density > 10,000 p km^{-2}.
LOWDENS.LST	Location of 217 areas with population density < one p km^{-2}.
CITY.LST	Population, Name, Latitude, Longitude of 2,763 cities
DRAWDOTS.BAS	Program to draw map with centroids.
PLOTCENT.BAS	Program to draw centroids on an equal area map.
PLOTCITY.BAS	Program to draw world map with city locations.
WORLDOUT.LST	World Data Bank Zero (Hershey 1965) - Coastal outlines.
GRENLAND	Town locations and estimated populations.
TASMANIA	Directory with sample files.



Global Demography Project
NCGIA/Geography, UC Santa Barbara


Global Demography Project
NCGIA/Geography, UC Santa Barbara


Global Demography Project
NCGIA/Geography, UC Santa Barbara

North and Latin America Medium Resolution Population Database   Data Summary Sheet

Num	Abb	UN Code	Country	Level used	Num of units	Ref year	Source	Country population (in 000) in ref yr (source)	in 1994 (proj.)	in 1994 (UN est)	Abs % diff.	Annual rates of change 80-85	85-90	90-95	Mean resolution in km	Mean pop per unit (in 000)	Area in sqkm
1	AI	660	Anguilla	0	1	92	2	8	9	8	15.10	0.60	1.42	1.32	9.5	9	91
2	AG	28	Antigua & Barbuda	0	1	91	2	64	65	67	2.67	0.89	0.68	0.71	21.0	65	442
3	AR	32	Argentina	1	23	91	19	32,616	33,797	33,875	0.23	1.43	1.27	1.17	347.5	1469	2,777,815
4	AW	533	Aruba	0	1	91	2	67	67	62	8.18	0.26	0.08	0.25	13.9	67	193
5	BS	44	Bahamas	0	1	90	2	255	272	272	0.08	2.07	1.82	1.60	117.7	272	13,865
6	BB	52	Barbados	0	1	90	2	257	261	260	0.24	0.29	0.35	0.33	20.7	261	430
7	BZ	84	Belize	1	6	85	8	166	208	206	0.77	2.63	2.60	2.03	61.9	35	22,965
8	BD	60	Bermuda	0	1	93	2	60	60	63	4.80	0.72	1.70	0.71	7.3	60	54
9	BO	68	Bolivia	1	9	88	14	6,405	7,384	7,888	6.39	2.55	2.46	2.37	349.4	820	1,098,575
10	BR	76	Brazil	2	2278	91	18	146,155	151,897	159,000	4.47	2.23	1.90	1.59	61.1	67	8,515,065
11	VG	92	British Virgin Islands	0	1	91	2	17	18	18	1.08	3.21	3.21	2.88	12.4	18	153
12	CA	124	Canada	2	289	91	25	27,292	28,402	28,147	0.91	0.90	1.13	1.38	185.3	98	9,922,385
13	KY	136	Cayman Islands	0	1	89	2	25	32	32	0.70	4.04	4.74	4.20	16.1	32	259
14	CL	152	Chile	1	13	92	17	13,348	13,773	14,026	1.81	1.68	1.66	1.55	240.5	1059	751,625
15	CO	170	Colombia	1	31	93	2	33,951	34,415	34,545	0.38	2.11	1.83	1.66	191.7	1110	1,138,915
16	CR	188	Costa Rica	2	81	92	6	3,160	3,319	3,347	0.82	2.91	2.77	2.41	25.1	41	50,900
17	CU	192	Cuba	2	169	91	21	10,793	11,102	11,000	0.93	0.81	1.03	0.89	26.0	66	114,525
18	DM	212	Dominica	0	1	91	2	71	71	71	0.46	-0.38	-0.38	-0.19	27.4	71	751
19	DO	214	Dominican Republic	0	30	90	23	7,168	7,759	7,769	0.13	2.38	2.22	1.98	40.2	259	48,440
20	EC	218	Ecuador	1	20	90	12	9,598	10,542	11,566	8.86	2.72	2.50	2.28	151.9	527	461,475
21	SV	222	El Salvador	1	14	90	9	5,272	5,752	5,641	1.97	0.92	1.75	2.18	39.1	411	21,395
22	FK	238	Falkland Islands	0	1	91	2	2	2	2	6.80	0.62	0.19	0.30	110.3	2	12,175
23	GF	254	French Guiana	0	1	90	2	114	130	141	7.65	3.69	3.39	3.04	301.7	130	91,000
24	GL	304	Greenland	0	1	93	-	55	58	58		1.15	1.04	0.71	1,475.0	58	2,175,600
25	GD	308	Grenada	0	1	91	2	95	96	92	3.92	0.20	0.20	0.31	18.6	96	345
26	GP	312	Guadeloupe	0	1	90	2	387	411	410	0.16	1.65	1.86	1.22	42.2	411	1,780
27	GT	320	Guatemala	2	326	94	3	10,322	10,322	10,322	0.00	2.82	2.88	2.88	18.3	32	108,890
28	GY	328	Guyana	1	3	91	2 / 20	740	755	825	8.49	0.80	0.15	0.94	267.7	252	214,970
29	HT	332	Haiti	0	1	92	2	6,764	7,045	7,035	0.14	1.83	2.01	2.03	166.6	7045	27,750
30	HN	340	Honduras	2	286	88	4	4,463	5,367	5,797	7.42	3.59	3.18	3.00	19.8	19	112,085
31	JM	388	Jamaica	1	15	91	22	2,243	2,408	2,521	4.50	1.60	0.92	1.02	27.6	161	11,425
32	MQ	474	Martinique	0	1	90	2	340	375	374	0.15	0.86	1.13	0.92	32.8	375	1,079
33	MX	484	Mexico	2	2403	90	10	81,855	92,381	91,840	0.59	2.40	2.22	2.06	28.7	38	1,972,545

34 MS	500	Montserrat	0	1	87	2	12	12	11	4.65	-0.55	-0.55	-0.34	10.2	12	104
35 AN	530	Netherlands Antilles	0	1	91	2	191	192	176	8.85	0.54	-0.02	0.12	28.3	192	800
36 NI	558	Nicaragua	1	16	82	5	3,384	4,275	4,275	0.00	2.83	2.60	3.74	96.2	267	148,000
37 PA	591	Panama	2	66	90	7	2,369	2,562	2,611	1.87	2.17	2.07	1.90	34.5	39	78,515
38 PY	600	Paraguay	1	20	91	15	4,397	4,773	4,767	0.14	3.20	2.93	2.69	142.6	239	406,750
39 PE	604	Peru	1	24	94	13	24,496	24,496	23,381	4.77	2.31	2.08	2.03	231.4	1021	1,285,215
40 PR	630	Puerto Rico	0	1	90	2	3,522	3,648	3,658	0.28	1.04	0.89	0.89	94.7	3648	8,960
41 KN	659	Saint Kitts & Nevis	0	1	88	2	44	43	41	4.65	-0.51	-0.51	-0.30	16.2	43	261
42 LC	662	Saint Lucia	0	1	91	2	136	142	141	0.53	1.42	1.42	1.35	24.8	142	616
43 VC	670	St. Vincent	0	1	91	2	108	111	111	0.00	0.86	0.86	0.88	19.7	111	389
44 SR	740	Suriname	1	9	91	1 / 2	404	428	455	5.93	1.66	1.94	1.86	134.9	48	163,820
45 TT	780	Trinidad & Tobago	0	1	94	2	1,292	1,292	1,292	0.00	1.40	1.28	1.08	71.6	1292	5,130
46 TC	796	Turks & Caicos Islar	0	1	90	2	12	15	14	3.66	4.33	4.33	3.85	20.7	15	430
47 US	840	United States	2	3141	90	24/26	247,606	258,833	260,513	0.64	0.92	0.94	1.03	54.6	82	9,363,130
48 UY	858	Uruguay	1	19	85	16	2,931	3,085	3,167	2.60	0.64	0.56	0.58	99.2	162	186,925
49 VI	850	U.S. Virgin Islands	1	3	90	2	102	102	108	5.91	1.82	0.01	0.15	10.7	34	345
50 VE	862	Venezuela	1	23	90	11	18,191	19,858	21,051	5.67	2.66	2.37	2.12	199.1	863	912,045
				sum	**avg**		**sum**	**sum**	**sum**	**avg**	**avg**	**avg**	**avg**	**avg**	**avg**	**sum**
				9341	**90**		**713,325**	**752,418**	**763,052**	**3.08**	**1.66**	**1.58**	**1.51**	**115.3**	**473**	**42,231,397**

Notes By Country:

15 CU	Includes urban divisions of Havana
19 SV	adjusted for "saldo migratorio"
24 GT	1981 census totals uniformly adjusted to match 1994 estimates because of significant underenumeration in 81 census
25 GY	1991 provincial populations derived from 1970 provincial populations and 1991 national population.
33 NI	1982 census totals uniformly adjusted to match 1994 UN estimates

POPULATION INFORMATION SOURCES

AMERICAS

1) Munro, D., 1990, Cambridge World Gazetteer: A Geographical Dictionary.
New York, Cambridge University Press.

2) The Europa World Yearbook 1994, 1994, London, Europa Publications Limited.

3) Censos Nacionales IX Censo de Poblacion 1981 de Guatemala
Republica de Guatemala, Instituto Nacional de Estadistica, Guatemala: 1984

4) Poblacion Total y Numero de Viviendas por Departamento y Municipio
Censo Nacional de Poblacion y Vivienda 1988, resultados definitivos
Tegucigalpa, D.C., Honduras: 1989

5) Landerbericht: Nicaragua 1988, C.E. Poeschel Verlag, et al., Federal Statistical Office, Wiesbaden: 1988

6) Costa Rica, Estimaciones y Proyecciones de Poblacion Cantonal.
Ministerio de Planificacaion Nacional y Politica Economica
Direccion General de Estadistica y Censos, San Jose, 1990.

7) Censos Nacionales de Poblacion y Vivienda 1990 de Panama, Republica de Panama,
Direccion de Estadistica y Censo, Panama City, Panama: 1991-1992

8) Landerbericht: Belize 1989, J.B. Metzler & C.E. Poeschel Verlag,
Federal Statistical Office, Wiesbaden: 1989

9) El Salvador, Official 1991 Estimates, Direccion General de Estadistica y Censos
Avance Estadistico, No. 19, San Salvador, June 1994

10) Market Mexico
INEGI XI General Census of Population and Housing 1990
Demosphere International, Inc., Fairfax, VA: 1993

11) El Censo 90 en Venezuela, Resultados Basicos
Central de Estadistica e Informatica, OCEI, Caracas, 1993.

12) V Censo de Poblacion y IV de Vivienda 1990 del Ecuador.
Republica del Ecuador, Instituto Nacional de Estadistica y Censos, Quito, 1990.

13) 1994 Estimates based on 1981 Census, Republica del Peru
Consejo Nacional de Poblacion, Peru: Hechos y Cifras Demograficas, Lima, 1989.

14) Encuesta Nacional de Poblacion y Vivienda 1988 de Bolivia
Resultados Finales, Republica de Bolivia, Instituto Nacional de Estadistica, La Paz, 1989.

15) Paraguay, Official 1991 Estimates, Direccion General de Estadistica,
Anuario Estadistico del Paraguay 1991, Asuncion, 1993.

16) Censo General de Poblacion y de Viviendas 1985 de Uruguay
Republica de Uruguay, Direccion General de Esatdistica y Censos, Montevideo, 1989.

17) Censo Nacional de Poblacion de Chile 1992
Republica de Chile, Instituto Nacional de Estadisticas, Santiago, 1992.

18) Censo Demografico 1991: resultados preliminares.
Fundacao Instituto Brasileiro de Geografia e Estatistica, Rio de Janeiro, 1992

19) Censo Nacional de Poblacion y Vivienda 1991 de Argentina
Republica Argentina, Instituto Nacional de Esadistica y Censos, Buenos Aires, 1992.

20) Landerbericht: Guyana 1987, C.E. Poeschel, et al., Federal Statistical Office, Wiesbaden: 1987

21) Cuba: Census of Cuba 1991 - unpublished document obtained from UN Statistical Division

22) Jamaica Population Census 1991, Population Census Office, Statistical Institute of Jamaica, Kingston, 1991.

23) Landerbericht: Donimikanische Republik 1990, J.B. Metzler/C.E Poeschel, Federal Statistical Office, Wiesbaden, 1990

24) ArcUSA - ESRI, Redlands

25) Statistics Canada: Census Divisions, Digital Boundaries and Data,
1991 Census of Canada, Cat. No. 93-304D, Ottawa: Statistics Canada, 1992

26) US Department of Commerce, Economics and Statistics Administration
Bureau of the Census, 1990 Census of Population and Housing - Alaska, March, 1993

BOUNDARY INFORMATION SOURCES

AMERICAS

For information on how to obtain any of the data, please contact the address listed under "Contact"

Canada

Canadian Census Divisions (CDs) have been purchased from Statistics Canada. Boundaries and pop data had been delivered separately and merged within the coverage. Data received in lat/long coordinates.
2nd level sub-national boundaries. Base map scale unknown.
Contact:
Statistics Canada, Licensing Services, Marketing and Information Services Branch
Ottawa, Ontario, Canada, K1A 0T6
Tel: 613-951-8211, Fax: 613-951-1134

Mexico

The coverage was obtained from Demosphere International, Inc. and merged into the Central American coverage. Data received in lat/long coordinates. 2nd level sub-national boundaries (municipios). Base map scale unknown.
Contact:
Demosphere International
4300 Fair Lakes Court, Suite 300
Fairfax, VA 22033, USA

Tel: 703-802-0100, Fax: 703-802-0102

United States

Boundaries and population data taken from Arc/USA. Projection has been changed to lat/long from Albers' conical equal area projection (standard parallels 25.5 and 45.5 degrees North, origin 96 degrees West and 23 degrees North). 2nd level sub-national boundaries (counties), except for Alaska which contains no county level data.
Contact:
ESRI, Inc.
380 New York St.
Redlands, CA 92373-8118
Tel: 909-793-2853, Fax: 909-793-5953

Guatemala, Belize, Nicaragua, El Salvador, Costa Rica, Panama, Bolivia, Brazil, Colombia, Ecuador, Paraguay, Peru, Uruguay, Venezuela

The coverages were obtained from the Centro Internacional de Agricultura Tropical, GIS Unit. The individual national coverages were merged into a Central America and a South America coverage respectively and population figures were added to these coverages. Data received in lat/long coordinates.
2nd level sub-national boundaries (municipios) for most of Central America and 18 states in Brazil. First subnational boundaries for the rest of South America. Base map scale unknown.
Contact:
Centro Internacional de Agricultura Tropical (CIAT)
GIS Unit
P.O. Box 6713
Cali, Colombia
Tel: 57 23 67 5050, Fax: 57 23 64 7243

Cuba

Data produced by the National Statistical Office of Cuba in collaboration with the UN Statistical Division's Software development project. Data were obtained in POPMAP format, converted to Arc/Info format and rubber sheeted into geographic coordinates. Coverage was subsequently merged into the Central America coverage. Original projection and map scale unknown. 2nd level sub-national boundaries.
Contact:
United Nations Statistical Division, Software Development Project
2, United Nations Plaza, New York, NY 10017
Fax: 212-963-1940

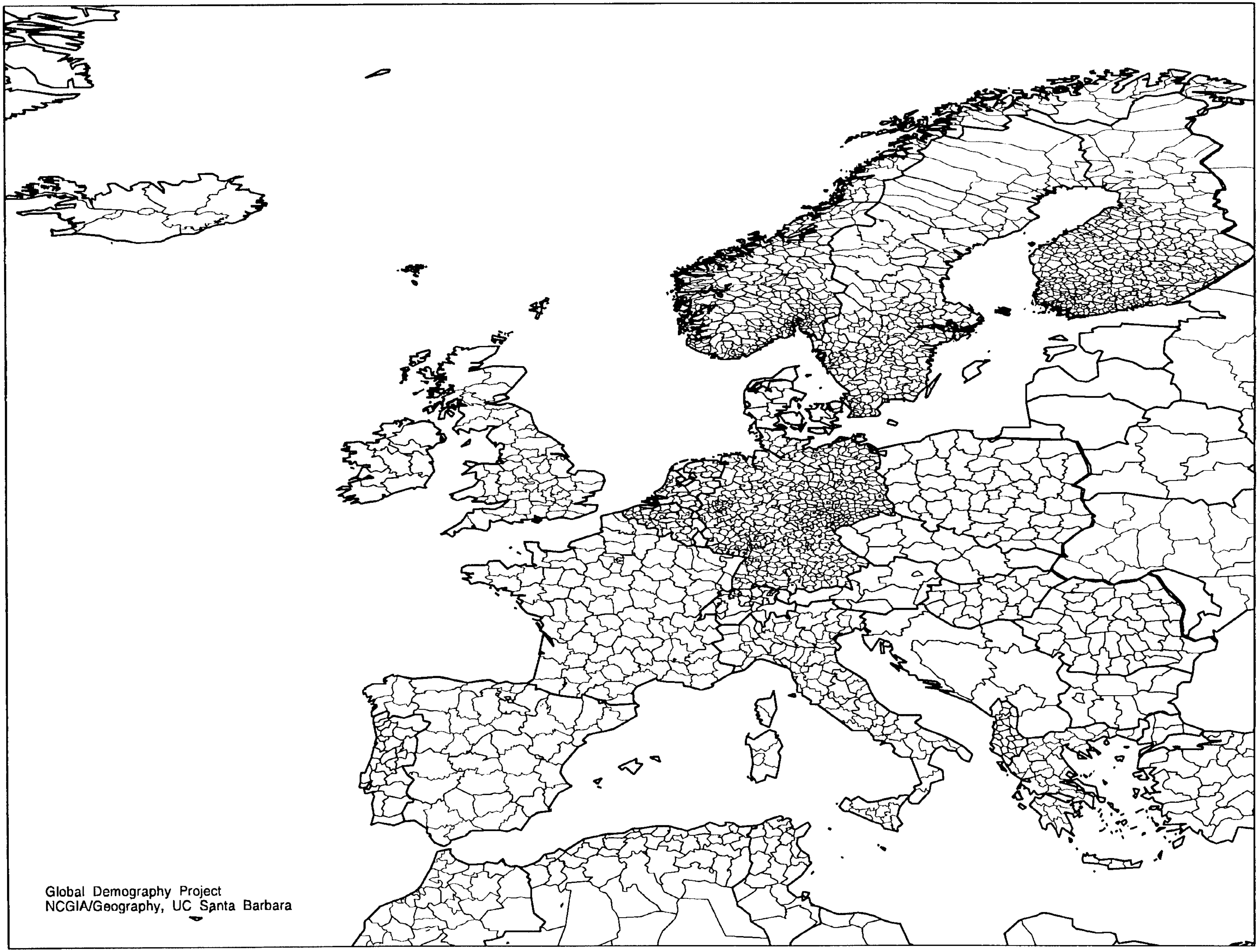
Dominican Republic

Boundaries were digitized at NCGIA from a 1:600,000 source map published by the Instituto Geografico Universitario, Universidad Autonoma de Santo Domingo, 1990. Map was subsequently transformed into lat/long. Map projection not published. 1st level sub-national boundaries.
Contact:

NCGIA
Geography Department
University of California, Santa Barbara, CA 93106-4060

Remaining Caribbean & South America

For those countries for which no higher resolution data were available, the boundaries were taken from ArcWorld which is based on the World Boundary Data Bank II and published by ESRI. Uses Latitude/longitude coordinates. 1st level for South America; mostly national level only for Caribbean Islands. 1:3 million nominal map scale.
Contact: See above (USA source)


Global Demography Project
NCGIA/Geography, UC Santa Barbara

Europe Medium Resolution Population Database **Data Summary Sheet**

Num	Abb	UN Code	Country	Level used	Num of units	Ref year	Source	Country population (in 000) in ref yr (source)	Country population (in 000) in 1994 (proj.)	Country population (in 000) in 1994 (UN est)	Abs % diff.	Annual rates of change 80-85	Annual rates of change 85-90	Annual rates of change 90-95	Mean resolution in km	Mean pop per unit (in 000)	Area in sqkm
1	AL	8	Albania	1	26	90	8	3,256	3,417	3,362	1.63	2.07	1.85	0.84	33.3	131	28,750
2	AN	20	Andorra	0	1	90	2	54	55	48	15.28	4.58	1.05	0.42	21.6	55	465
3	AU	40	Austria	1	9	91	3	7,660	7,755	7,834	1.00	0.02	0.40	0.38	96.5	862	83,855
4	BE	56	Belgium	2	43	89	5	5	10,032	10,020	0.12	0.01	0.22	0.13	26.6	233	30,520
5	BU	100	Bulgaria	1	9	90	2	8,989	8,943	8,903	0.45	0.22	0.07	-0.23	111.0	994	110,910
6	CL	830	Channel Islands	0	1	89	2	83	86	144	40.03	0.54	1.35	0.24	13.9	86	194
7	CZ	203	Czechoslovakia	1	10	84	1	13,849	13,980	15,818	11.62	0.25	0.21	0.27	113.1	1,398	127,870
8	DA	208	Denmark	1	14	89	5	5	4,668	5,180	9.89	0.75	0.60	0.20	55.5	333	43,075
9	EN	233	Estonia	0	1	90	12	1,583	1,591	1,574	1.07	0.75	0.60	-0.15	212.4	1,591	45,100
10	FO	234	Faeroe Islands	0	1	92	2	47	47	47	0.14	0.95	0.62	0.24	37.4	47	1,399
11	FI	246	Finland	2	458	91	11	4,995	5,031	5,033	0.03	0.51	0.32	0.26	27.1	11	337,030
12	FR	250	France	2	96	89	5	56,436	57,757	57,571	0.32	0.47	0.55	0.37	75.3	602	543,965
13	GM	276	Germany	2	543	91	2 & 5	80,274	81,463	80,942	0.64	-0.16	0.46	0.44	25.6	150	356,840
14	GI	292	Gibraltar	0	1	91	2	28	29	31	7.28	-0.61	1.62	0.26	2.6	29	7
15	GR	300	Greece	2	51	89	5	10,154	10,307	10,232	0.74	0.60	0.38	0.26	50.9	202	131,985
16	HU	348	Hungary	1	19	90	7	10,382	10,310	10,479	1.61	-0.12	-0.18	-0.16	70.0	543	93,030
17	IC	352	Iceland	1	7	80	13	229	267	266	0.47	1.13	1.08	1.04	121.2	38	102,820
18	EI	372	Ireland	1	10	89	5	5,098	5,016	3,476	44.30	0.87	-0.28	-0.19	83.0	502	68,895
19	IM	833	Isle of Man	0	1	91	2	70	71	69	3.33	-0.15	1.58	0.24	24.0	71	576
20	IT	380	Italy	2	95	89	5	57,541	57,909	57,869	0.07	0.25	0.18	0.09	56.3	610	301,245
21	LV	428	Latvia	0	1	90	12	2,687	2,690	2,658	1.21	0.61	0.55	-0.28	252.4	2,690	63,700
22	LS	438	Liechtenstein	0	1	90	2	28	29	28	4.79	1.04	0.82	0.17	12.6	29	160
23	LH	440	Lithuania	0	1	90	12	3,723	3,787	3,765	0.57	0.88	0.79	0.21	255.3	3,787	65,200
24	LU	442	Luxembourg	0	1	89	5	377	387	383	1.06	0.15	0.35	0.69	50.8	387	2,585
25	MT	470	Malta	0	1	91	2	360	366	364	0.66	1.21	0.55	0.72	17.8	366	316
26	MN	492	Monaco	0	1	90	2	27	27	28	2.11	0.75	0.35	0.35	1.4	27	2
27	NL	528	Netherlands	2	39	89	5	14,847	15,447	15,384	0.41	0.48	0.62	0.93	32.5	396	41,160
28	NO	578	Norway	2	436	91	9	4,264	4,329	4,333	0.10	0.33	0.45	0.51	27.3	10	323,895
29	PL	616	Poland	1	48	91	2	37,543	37,912	38,621	1.84	0.90	0.52	0.29	80.7	790	312,685
30	PO	620	Portugal	1	28	89	5	9,391	9,369	9,877	5.14	0.28	-0.07	0.03	57.2	335	91,630

31	RO	642	Romania	1	41	89	6	23,152	23,541	23,434	0.45	0.47	0.42	0.26	76.1	574	237,500
32	SM	674	San Marino	1	1	92	2	23	24	23	3.30	0.86	0.55	0.02	7.8	24	61
33	SP	724	Spain	2	50	89	5	37,411	37,757	39,214	3.72	0.49	0.25	0.16	100.5	755	504,880
34	SW	752	Sweden	2	285	90	10	8,547	8,728	8,732	0.04	0.10	0.51	0.48	39.7	31	449,790
35	SZ	756	Switzerland	1	18	90	4	5,654	6,714	6,910	2.84	0.47	0.93	0.71	47.9	373	41,285
36	UK	826	United Kingdom	2	64	89	5	55,654	56,420	57,956	2.65	0.10	0.28	0.24	61.8	882	244,755
37	YO	891	Yugoslavia	1	8	81	1	26,044	23,720	24,049	1.37	0.72	0.58	0.25	178.8	2,965	255,805
					sum	**avg**		**sum**	**sum**	**sum**	**avg**	**avg**	**avg**	**avg**	**avg**	**avg**	**sum**
					2,420	**89**		**490,470**	**509,985**	**514,657**	**4.66**	**0.61**	**0.57**	**0.29**	**69.1**	**619**	**5,043,940**

Notes By Country:

13 GM Data for former West Germany is for 1989 (from EUROSTAT), for former East Germany, the reference year is 1991 (Statistical Yearbook of 1993)

11 FI Seven second level units for which no data existed in census publication.

18 IC First level units Reykjavik and Reykjaneskjord aggregated (Reykjaneskjord) due to coverage boundary problems

26 MT Excl. Non-Maltese population

POPULATION INFORMATION SOURCES

EUROPE

1) Munro, D., 1990, Cambridge World Gazetteer: A Geographical Dictionary, New York, Cambridge University Press.

2) The Europa World Yearbook 1994, London, Europa Publications Limited.

3) Hauser- und Wohnungszahlung 1991: Hauptergebnisse Osterreich. Beitrage zur Osterreichischen Statistik, Wien, 1993

4) Annuaire statistique de la Suisse 1993, (Statistisches Jahrbuch der Schweiz, Bundesamt fur Statistik), Office federal de la statistique, Zurich, 1992.

5) Eurostat data contained in digital spatial data set

6) Anuarul Statistic al Romaniei 1990 Comisia Nationala Pentru Statistica, Romania, 1990.

7) 1990 Population census, Department of Population Surveys, Hungarian Central Statistical Office, Budapest, 1993

8) Landerbericht: Albanien 1993, J.B. Metzler & C.E. Poeschel Verlag, Federal Statistical Office, Wiesbaden.

9) Statistical Yearbook of Norway 1992 Central Bureau of Statistics of Norway, Oslo-Kongsvinger, 1992

10) Statistical Abstract of Sweden 1991 Statistiska Centralbyran, Stockholm, 1991

11) 1991 Population Census, Statistics Finland, 1991.

12) Baseline Demographics for the Republics of the Former Soviet Union New World Demographics, (see below), 1992.

13) Population and Vital Statistics, Iceland, Statistical Bureau of Iceland, Reykjavik, 1988

14) Petit Manual Statistique, Office Federal de la Statistique, Beograd, 1992.

15) Statisticka Rocenska Ceske Republiky, Cesky Statisticky Urad, Praha, 1993.

BOUNDARY INFORMATION SOURCES

EUROPE

For information on how to obtain any of the data, please contact the address listed under "Contact"

Eastern & Western Europe

Boundary and population data were obtained from the Statistical Office of the European Communities (EUROSTAT). Data were converted to Arc/Info format and projected into geographic (lat/long) coordinate from the Lambert Azimuthal equal area projection. First to third subnational level, depending on country. Base map scale unknown.
Contact:

Statistical Office of the European Communities
GISCO
L-2820 Luxembourg
Tel: 352 4301 34398, Fax: 352 4301 32594

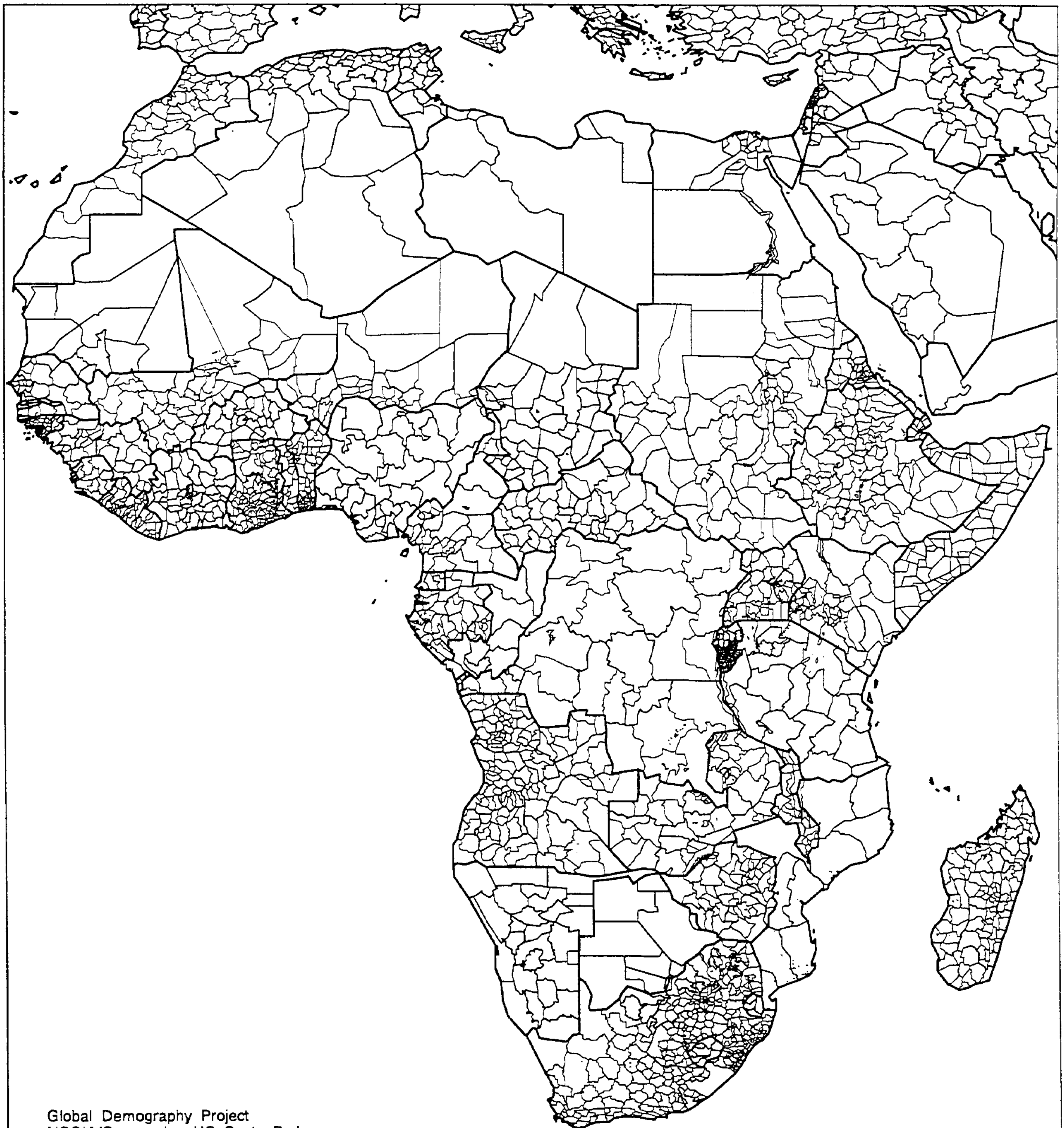
Community of Independent States (former Soviet Union)

Data and boundaries were obtained from New World Demographics in Arc/Info format. Further processing was required to remove artificial polygon splits that were due to PC GIS software limitations. Projected into latitude/longitude coordinates from Lambert Conformal projection (standard parallels at 47 and 62 degrees north latitude and central meridian at 90 degrees east longitude). Oblast level - formerly 2nd subnational level in the Soviet Union. Base map scale 1:10 million.
Contact:

New World Demographics
P.O. Box 866
Shady Side, MD 20764-0866, USA
Tel: 410-867-3767
Fax: 410-867-0549

ADDENDUM

Several regions of the former Yugoslavia are still listed under that name in the materials accompanying this report. Although Croatia and Slovenia are now independent countries, and the Former Yugoslav Region of Macedonia is known by that name (FYROM), complete population data are available only for the period previous to the dissolution of the former Yugoslavia. As well, the designations, and perhaps boundaries, for the other parts of this area are still not final. Fortunately, these changes have no consequence for the population rasters, but they do bring the country count to 219.



Global Demography Project
NCGIA/Geography, UC Santa Barbara

Africa Medium Resolution Population Database Data Summary Sheet

Num	Abb	UN Code	Country	Level used	Num of units	Ref year	Source	Country population (in 000) in ref yr (source)	in 1994 (proj.)	in 1994 (UN est)	Abs. % diff.	Annual rates of change 80-85	85-90	90-95	Mean resolution in km	Mean pop per unit (in 000)	Area in sqkm
1	AG	4	Algeria	1	31	84	1	20,841	27,459	27,815	1.28	3.07	2.72	2.71	274	886	2,321,829
2	AO	7	Angola	1	18	91	27	10,310	11,527	10,674	7.99	2.63	2.84	3.72	263	640	1,245,530
3	BN	53	Benin	1	6	79	28	3,331	5,175	5,235	1.14	2.82	2.98	3.11	139	863	116,267
4	BC	20	Botswana	1	10	91	1	1,325	1,447	1,392	3.92	3.40	2.95	2.92	241	145	580,015
5	UV	233	Burkina Faso	1	29	85	29	7,965	10,165	10,069	0.95	2.50	2.64	2.81	97	351	273,735
6	BY	29	Burundi	2	114	90	22	5,357	6,011	6,168	2.54	2.80	2.90	2.88	15	53	25,326
7	CM	32	Cameroon	2	31	87	30	10,822	13,218	12,905	2.43	2.83	2.90	2.83	122	426	462,564
8	SP	203	Canary Islands	1	3	80	5	1,445	1,511	1,511	0.00	0.49	0.25	0.16	50	504	7,636
9	CV	35	Cape Verde	1	10	80	5	296	414	407	1.61	2.26	2.27	2.88	21	41	4,288
10	CT	37	Central African Rep.	2	48	88	32	2,688	3,149	3,344	5.82	2.58	2.68	2.62	114	66	621,503
11	CD	39	Chad	1	14	88	6	5,428	6,309	6,183	2.03	2.28	2.03	2.71	301	451	1,266,188
12	CN	45	Comores Islands	1	4	80	2	394	635	630	0.74	3.42	3.56	3.68	23	159	2,030
13	CF	46	Congo	1	9	83	1	1,675	2,318	2,515	7.82	2.82	2.97	3.00	196	258	344,994
14	IV	107	Cote d'Ivoire	1	50	88	18	10,813	13,499	13,895	2.85	3.86	3.74	3.68	80	270	320,000
15	DJ	72	Djibouti	1	5	83	2	259	451	496	9.12	4.46	2.92	3.01	65	90	21,356
16	EG	59	Egypt	1	26	86	1	46,718	56,133	57,285	2.01	2.58	2.39	2.20	194	2159	977,956
17	EK	61	Equatorial Guinea	1	6	89	35	341	386	389	0.68	7.22	2.43	2.55	67	64	27,206
18	ER	997	Eritrea	2	38	84	19	2,747	3,662	3,437	6.55	2.12	2.91	3.05	56	96	120,311
19	ET	62	Ethiopia	3	550	84	19	39,868	53,143	53,435	0.55	2.12	2.91	3.05	45	97	1,125,789
20	GB	74	Gabon	2	33	90	36	1,365	1,561	1,323	18.00	4.01	3.25	3.31	89	47	259,806
21	GA	75	Gambia, The	1	5	83	2	688	936	956	2.09	3.01	2.89	2.60	46	187	10,632
22	GH	81	Ghana	2	107	84	43	12,206	16,698	16,944	1.45	3.58	3.14	3.00	46	156	227,202
23	GV	90	Guinea	1	7	80	3	4,300	6,242	6,501	3.98	2.23	2.86	3.04	187	892	245,157
24	PU	175	Guinea-Bissau	1	9	79	2	779	1,086	1,050	3.41	1.87	1.99	2.14	61	121	33,101
25	KE	114	Kenya	2	43	89	8	21,836	25,835	26,975	4.23	3.56	3.42	3.35	116	601	577,038
26	LT	122	Lesotho	1	8	86	13	1,578	1,928	1,929	0.04	2.79	2.54	2.47	62	241	30,352
27	LI	123	Liberia	1	9	84	37	2,102	2,902	2,941	1.31	3.17	3.16	3.32	104	322	96,826
28	LY	124	Libya	1	7	84	1	3,637	5,246	5,225	0.39	4.37	3.65	3.47	481	749	1,618,295
29	MA	129	Madagascar	2	93	90	6/42	11,443	13,047	13,702	4.78	3.06	3.18	3.29	80	140	591,882
30	PO	998	Madeira	1	1	81	5	254	256	256	0.00	0.28	-0.07	0.03	29	256	870
31	MI	130	Malawi	2	24	87	10	7,989	10,660	11,008	3.16	3.42	5.34	3.31	63	444	94,455
32	ML	133	Mali	2	41	87	4/11	7,834	9,745	10,464	6.87	2.85	3.04	3.17	175	238	1,255,090
33	MR	136	Mauritania	1	12	92	38	2,081	2,204	2,270	2.90	2.60	2.73	2.86	295	184	1,041,970

34	MP	137	Mauritius	1	1	90	39	1,037	1,097	1,109	1.06	1.09	1.05	1.00	36	1097	1,328
35	MO	143	Morocco	2	42	90	1	25,208	27,768	27,601	0.60	2.56	2.58	2.40	127	661	674,167
36	MZ	144	Mozambique	1	10	87	1	14,361	16,605	15,823	4.94	2.27	0.94	2.83	279	1660	779,655
37	WA	147	Namibia	1	26	91	44	1,410	1,551	1,635	5.14	2.94	3.06	3.18	177	60	818,346
38	NG	158	Niger	2	35	88	40	7,250	8,798	8,813	0.17	3.36	3.14	3.26	184	251	1,182,868
39	NI	159	Nigeria	1	31	91	1	88,515	97,229	123,079	21.00	3.20	3.30	3.13	171	3136	903,115
40	RE	182	Reunion	1	1	94	7	644	644	644	0.00	1.65	1.87	1.56	45	644	2,036
41	RW	184	Rwanda	1	10	91	23	7,165	7,934	8,057	1.52	3.27	2.87	3.40	48	793	22,950
42	TP	193	Sao Tome & Princ.	1	2	83	5	100	129	130	1.07	2.39	2.32	2.20	21	64	856
43	SG	195	Senegal	1	10	88	21	6,893	8,117	8,165	0.59	2.81	2.78	2.70	140	812	196,820
44	SE	999	Seychelles	1	1	94	7	73	73	73	0.00	2.87	3.29	0.89	14	73	210
45	SL	197	Sierra Leone	2	12	85	12	3,516	4,552	4,616	1.39	2.93	3.07	2.66	77	379	71,609
46	SO	201	Somalia	1	16	80	26	6,713	9,952	9,845	1.08	3.19	1.94	3.18	200	622	640,061
47	SF	202	South Africa	2	292	91	25	37,015	43,349	41,749	3.83	2.58	2.44	2.37	64	148	1,214,484
48	SU	206	Sudan	2	85	83	17	20,181	27,713	28,175	1.64	3.11	2.88	2.78	172	326	2,505,048
49	WZ	209	Swaziland	1	4	86	16	681	843	836	0.81	3.08	2.62	2.68	66	211	17,165
50	TZ	215	Tanzania	1	23	88	9	23,174	28,386	29,755	4.60	3.28	3.43	3.36	196	1234	886,807
51	TO	217	Togo	2	21	81	15	2,720	4,048	4,010	0.96	2.93	3.07	3.18	52	193	57,281
52	TS	222	Tunisia	1	23	84	2	6,966	8,620	8,757	1.56	2.57	2.08	2.06	82	375	155,533
53	UG	226	Uganda	2	34	91	14	16,583	18,144	19,823	8.47	2.72	3.11	3.00	78	534	205,111
54	CG	250	Zaire	2	31	84	41	29,671	41,026	42,476	3.41	3.18	3.32	3.17	273	1323	2,314,389
55	ZA	251	Zambia	2	54	90	24	7,818	8,779	9,132	3.87	3.58	3.40	2.84	117	163	743,293
56	ZI	181	Zimbabwe	2	58	82	20	7,546	11,107	11,215	0.97	3.22	3.45	2.96	82	191	386,996
					sum	**avg**			**sum**	**sum**	**avg**	**avg**	**avg**	**avg**	**avg**	**avg**	**sum**
					2223	**86.2**			**691,423**	**724,291**	**3.13**	**2.89**	**2.75**	**2.75**	**123**	**485**	**724,291**

POPULATION INFORMATION SOURCES

AFRICA

1) Europa Publications, 1993, The Middle East and North Africa - 1993, 39th ed., London and (1992), Africa South of the Sahara - 1993, 22nd ed., London

2) U.S. Bureau of the Census (no date), International Database, Washington, D.C.

3) CIAT (1989), Secondary administrative units, Africa, Version 1, Centro Internacional de Agricultura Tropical, Cali, Colombia

4) ESRI (1984), UNEP/FAO African Database, Final Report, Environmental Systems Research Institute, Redlands

5) Munro, D., ed, 1988, Chambers World Gazetteer, 5th ed., Cambridge, Chambers.

6) Hunter, B., ed., 1993, The Statesman's Year-Book 1992-93, 129th edition, St. Martin's Press, New York

7) United Nations, 1993, World Population Prospects, The 1992 Revision, Department for Economic and Social, Information and Policy Analysis, New York

8) Republic of Kenya, 1989, Census of Kenya 1989, Nairobi

9) The United Republic of Tanzania, 1988. Sensa 1988 - 1988 Population Census: Preliminary Report, Bureau of Statistics, Ministry of Finance, Dar es Salaam

10) Malawi Government, 1991, Malawi Population and Housing Census 1987, Summary of Final Results, Vol 1, National Statistical Office, Zomba

11) Republique du Mali, 1991, Recensement General de la Population et de l'Habitat au Mali - 1987, Direction Nationale de la Statistique et de l'Information, Bamako

12) Sierra Leone Government, 1986, The preliminary report on the 1985 National Population Census of Sierra Leone, Freetown

13) Lesotho Government, 1987, 1986 Population Census - Preliminary Results, Bureau of Statistics, Maseru

14) Republic of Uganda, 1991, Provisional Results of the 1991 Population and Housing Census, Statistics Department, Ministry of Planning and Economic Development, Kampala

15) Republique Togolaise, 1986, Apercu des Resultats d'Ensemble du Recensement General de la Population et de l'Habitat au Togo 1981, Bureau Central du Recensement, Lome

16) Swaziland Government, 1986, Report on the 1986 Swaziland Population Census Vol 1, Central Statistical Office, Mbabane

17) University of Gezira, 1983, Population of the Sudan and its Regions, Population Studies Centre, Gezira

18) Republique de Cote d'Ivoire, 1988, Recensement General de la Population et de l'Habitat 1988. Resultats Provisoires, Comite National de Recensement, Abidjan

19) The boundaries and data files were produced by USGS and obtained from the Int'l Food Policy Research Institute, Washington D.C. via ILRAD, Nairobi. Additional data from:

Provisional Military Government of Socialist Ethiopia, 1988, Population of Weredas and towns by sex, and average household size ..., Central Statistical Office, Addis Ababa

20) Government of Zimbabwe, 1984, 1982 population census. A preliminary assessment, Central Statistical Office, Harare

21) Republique du Senegal, 1988, Les principaux resultats provisoires du recensement de la population de l'habitat du Senegal 1988, Direction de la statistique, Dakar

22) Republique du Burundi, 1990, Resultats Provisoires. Recensement General de la Population et de l'Habitat, Bureau Central de Recensement, Gitega

23) Republique Rwandaise, 1991, Recensement General de la Population et de l'Habitat au 15 Aout 1991. Resultats Provisoires, Service National de Recensement, Kigale

24) Republic of Zambia, 1990, 1990 Census of Population, Housing and Agriculture, Preliminary Report, Central Statistical Office, Lusaka

25) Republic of South Africa, 1991, Population Census 1991, Adjustment for Undercount, Central Statistical Service, Pretoria

Republic of South Africa, 1992, 1992 South African Statistics, Central Statistical Office, Pretoria (contains map of census regions)

Ciskei Central Statistical Service, 1988, Statistical Abstract of Ciskei 1988, No. 1.

26) Somali Democratic Republic, 1984, 1982 Statistical Abstract, Central Statistical Department, State Planning Commission, Mogadishu

27) Republica Popular de Angola, 1987, Boletim Demographico #4, Projeccao da Populacao do Pais por Provincias para o Period 1970/95, Instituto Nacional de Estatistica, Luanda

28) Republique Populaire du Benin, 1987, Recensement General de la Population et de l'Habitation (Mars 1979), Institut National de la Statistique, Cotonou

29) Republique Democratique Populaire de Burkina Faso, 1988, Annuaire Statistique 1988, Institute National de la Statistique et de la Demographie, Ouagadougou

30) Republique du Cameroun, 1988, Deuxieme Recensement General de la Population et de l'Habitat du Cameroun, Yaounde

31) Republica de Cabo Verde, 1989, Boletim Anual de Estatistica, Direccao General de Estatistica, Praia

32) Republique Centrafricaine, 1992, Annuaire Statistique 1991,
Ministere de l'Economie, du Plan, des Statistiques et de la Cooperation Internationale, Bangui

33) Republique du Djibouti, 1989, Annuaire Statistique de Djibouti,
Direction Nationale de la Statistique, Djibouti

34) Arab Republic of Egypt, 1991, Statistical Year Book,
Central Agency for Public Mobilisation and Statistics, Cairo

35) Republica de Guinea Equatorial, 1990, Boletin Estadistica
Anual 1990, Direccion General de Estadistica, Malabo

Republica de Guinea Equatorial, 1990, Censos Nacionales:
I de Poblacion y i de Vivienda de 4 de Julio de 1983,
Direccion General de Estadistica, Malabo

Republica de Guinea Equatorial (1987), Guinea en Cifras 1987,
Direccion General de Estadistica, Malabo

36) Republique Gabonaise, 1990, Annuaire Statistique du Gabon,
Direction Generale de la Statistique et des Etudes Economiques, Libreville

37) Government of Liberia, 1987, 1984 Population and Housing Census of Liberia - Summary Results
Ministry of Planning and Economic Affairs, Monrovia

38) Republique Islamique de Mauritanie, 1986, Annuaire Statistique,
Direction de la Statistique et de la Comptabilite Nationale, Nouakchott

39) Government of Mauritius, 1991, Annual Digest of Statistics 1990,
Central Statistical Office, Port Louis

40) Republique du Niger, 1989, 2eme Recensement General de la
Population 1988, Bureau Central du Recensement, Niamey

Republique du Niger, 1991, Annuaire Statistique << Series Longues>>,
Direction de la Statistique et de la Demographie, Niamey

41) Republique du Zaire, 1984, Combiens sommes-nous? Recensement
Scientifique de la Population, 1er Juillet 1984, Kinshasa

42) Republika Demokratika Malagasy, 1975, Recensement 1975 -
Analyse des Donnees Demographiques, Antananarivo

43) Republic of Ghana, 1984, 1984 Population Census of Ghana -
Preliminary Report, Central Bureau of Statistics, Accra

44) Republic of Namibia, 1993, Statistical Abstract 1993, Num. 2,
National Planning Commission, Center for Statistics, Windhoek

BOUNDARY INFORMATION SOURCES

AFRICA

For information on how to obtain any of the data, please contact the address listed under "Contact"

Ethiopia

Data were produced by the Famine Early Warning Systems (FEWS) project at EROS Data Center, Sioux Falls, and obtained via the International Livestock Research Institute in Nairobi. Data received in lat/long coordinates. 3rd level sub-national boundaries.
Contact:
FEWS Project
EROS Data Center
Sioux Falls, SD 57198
Tel: 605-594-6061

Gabon, Central African Republic, Benin

Data were digitized from IGN basemaps, in lat/long coordinates. 2nd level sub-national boundaries.
Contact: NCGIA (see above)

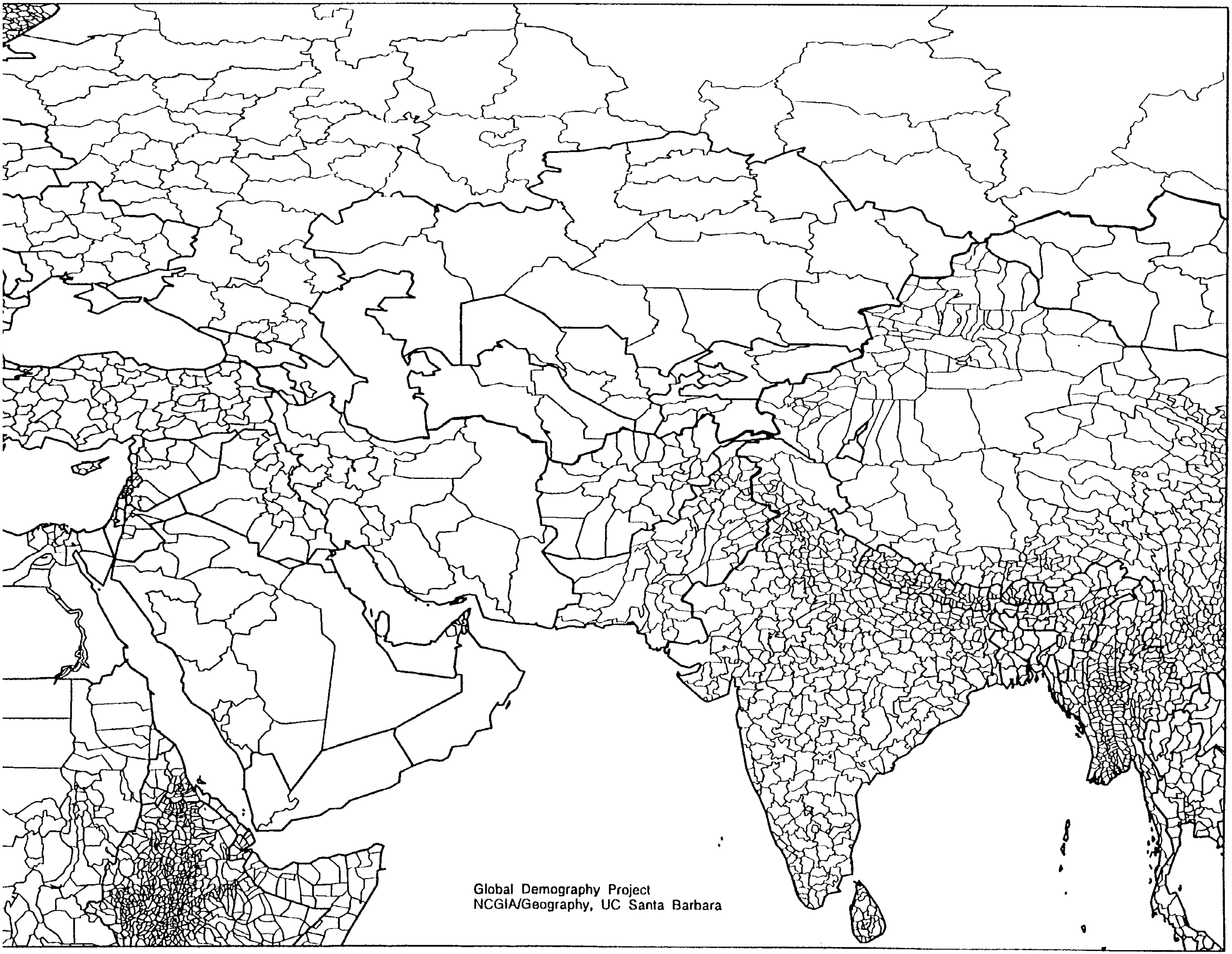
Zaire

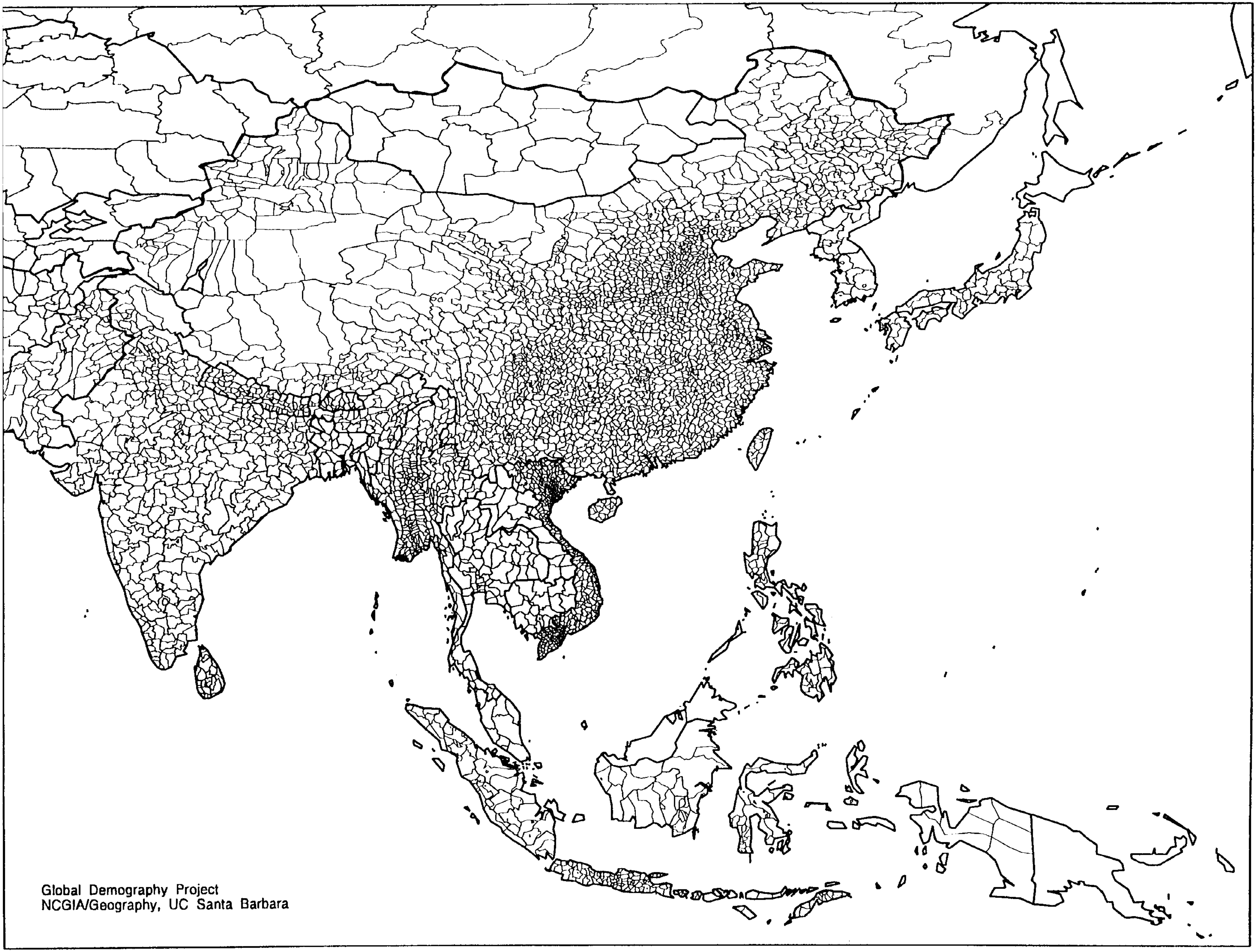
Data were obtained from UNESCO's Tropical Soils Fertility Programme. Data received in lat/long coordinates. 2nd level sub-national boundaries.
Contact:
UNESCO/TSBF
United Nations Avenue
P.O. Box 30592
Nairobi, Kenya

Remaining Africa

Data were produced by the UN Food and Agriculture Organization's GIS Unit partly based on the World Boundary Data Base II. All other Africa data sets were merged into this continental coverage. Data received in lat/long coordinates. 1st or 2nd level sub-national boundaries depending on country. Contact:

United Nations Food and Agricultural Organization
FAO Headquarters, GIS Center, Room B-113
Viale delle Terme di Caracalla
00100 Rome, Italy


Global Demography Project
NCGIA/Geography, UC Santa Barbara


Global Demography Project
NCGIA/Geography, UC Santa Barbara

Asia Medium Resolution Population Database   Data Summary Sheet

Num	Abb	UN Code	Country	Level used	Num of units	Ref year	Source	Country population (in 000) in ref yr (source)	in 1994 (proj.)	in 1994 (UN est)	Abs % diff.	Annual rates of change 80-85	85-90	90-95	Mean resolution in km	Mean pop per unit (in 000)	Area in sqkm
1	AF	4	Afghanistan	1	28	82	12	13,471	17,250	18,879	8.6	-2.02	2.63	6.74	152.62	616	652,225
2	AM	51	Armenia	0	1	90	21	3,293	3,377	3,548	4.8	0.89	0.84	0.51	173.21	3,377	30,000
3	AZ	31	Azerbaijan	1	3	90	21	5,351	7,472	7,472	0.0	0.89	0.84	0.51	170.29	2,491	87,000
4	BA	48	Bahrain	0	1	83	4	400	576	563	2.3	4.26	3.17	2.79	25.71	576	661
5	BG	50	Bangladesh	1	21	81	4	87,149	120,735	125,201	3.6	2.68	2.39	2.41	82.81	5,749	144,000
6	BO	112	Belarus	1	6	90	21	10,259	10,521	10,163	3.5	0.89	0.84	0.51	186.19	1,754	208,000
7	BT	64	Bhutan	1	18	85	5	1,286	1,586	1,689	6.1	2.10	2.19	2.33	50.89	88	46,620
8	BX	96	Brunei Darussalam	0	1	90	2	257	282	282	0.1	3.12	2.60	2.24	75.93	282	5,765
9	CB	116	Cambodia	1	19	81	2	7,022	9,130	9,222	1.0	2.40	2.58	2.50	97.60	481	181,000
10	CN	156	China	2	2422	87	20/24	115,767	1,275,203	1,222,017	4.4	1.44	1.49	1.42	62.95	527	9,597,000
11	CY	196	Cyprus	1	6	92	1 / 4	725	739	729	1.4	1.11	1.08	0.94	39.26	123	9,250
12	GZ	275	Gaza Strip	1	1	91	5	676	729	764	4.6	3.20	3.76	1.98	20.12	729	405.00
13	GO	268	Georgia	1	4	90	21	5,456	5,596	5,450	2.7	0.89	0.84	0.51	132.00	1,399	69,700
14	HK	344	Hong Kong	0	1	91	5	5,674	5,805	5,890	1.4	1.59	0.91	0.77	32.59	5,805	1,062
15	IN	356	India	2	453	91	5,9	836,952	894,609	913,747	2.1	2.14	1.97	1.91	83.61	1,975	3,166,830
16	ID	360	Indonesia	2	291	90	25	175,991	189,331	198,876	4.8	2.06	1.93	1.78	81.22	651	1,919,445
17	IR	364	Iran	1	24	86	6	49,445	64,193	64,878	1.1	4.40	3.50	2.71	262.04	2,675	1,648,000
18	IZ	368	Iraq	1	18	85	13	15,585	20,942	20,562	1.8	3.27	3.31	3.21	156.07	1,163	438,445
19	IS	376	Israel	1	6	83	4	4,038	5,695	5,672	0.4	1.75	1.92	4.67	58.84	949	20,770
20	JA	392	Japan	2	47	88	23	122,779	125,746	125,422	0.3	0.68	0.44	0.38	88.69	2,675	369,700
21	JO	400	Jordan	1	8	90	14	3,312	5,198	5,198	0.0	3.07	3.25	3.41	109.54	650	96,000
22	KA	398	Kazakhstan	1	17	90	21	16,691	17,117	17,027	0.5	0.89	0.84	0.51	399.80	1,007	2,717,300
23	KN	408	Korea, Dem. Rep.	1	12	87	18	19,318	22,035	23,491	6.2	1.71	1.81	1.88	100.96	1,836	122,310
24	KS	410	Korea, Republic	1	14	90	19	42,896	43,411	44,843	3.2	1.36	1.22	0.82	83.86	3,101	98,445
25	KU	414	Kuwait	0	1	80	4	1,693	1,693	1,693	0.0	4.48	4.40	-5.79	155.82	1,693	24,280
26	KI	417	Kyrgyzstan	1	3	90	21	4,367	4,479	4,667	4.0	0.89	0.84	0.51	257.23	1,493	198,500
27	LA	418	Laos	1	17	90	5	4,170	4,723	4,743	0.4	2.29	3.12	3.00	118.00	278	236,725
28	LE	422	Lebanon	0	1	92	5	2,838	2,943	2,965	0.7	-0.01	0.53	2.00	101.98	2,943	10,400
29	MC	446	Macau	0	1	91	5	356	389	398	2.4	3.90	3.33	2.86	4.12	389	17
30	MY	458	Malaysia	1	11	90	5	13,384	19,626	19,684	0.3	2.09	2.64	2.35	173.98	1,784	332,965
31	MV	462	Maldives	0	1	90	5	213	241	246	2.2	2.85	3.04	3.02	17.26	241	298
32	MD	--	Moldova	0	1	90	21	4,474	4,474	--	--	0.89	0.84	0.51	183.58	4,474	33,700

33	MN	496	Mongolia	1	18	90	5	2,003	2,228	2,363	5.7	2.76	2.74	2.63	294.86	124	1,565,000
34	BM	104	Myanmar	2	284	83	0	34,054	43,100	45,573	5.4	2.09	2.16	2.14	48.86	152	678,030
35	NP	524	Nepal	2	75	91	17	18,491	19,927	21,602	7.8	2.85	2.66	2.45	43.42	266	141,415
36	MU	512	Oman	0	1	86	26	2,017	2,090	2,077	0.6	4.91	3.76	3.57	521.49	2,090	271,950
37	PK	586	Pakistan	1	70	81	8	84,254	126,694	131,434	3.6	3.31	3.20	2.67	107.17	1,810	803,940
38	RP	608	Philippines	2	74	90	22	60,480	65,981	67,898	2.8	2.58	2.39	2.07	63.67	892	300,000
39	QA	634	Qatar	0	1	94	5	478	478	478	0.0	8.90	3.54	2.76	106.93	478	11,435
40	RU	643	Russia	2	71	90	21	148,041	151,828	147,370	3.0	0.89	0.84	0.51	490.44	2,138	17,078,005
41	SA	682	Saudi Arabia	0	1	92	5	16,929	18,100	17,035	6.3	5.57	3.67	3.38	1,549.48	18,100	2,400,900
42	SG	702	Singapore	0	1	90	5	2,705	2,824	2,826	0.1	1.15	1.15	1.03	24.82	2,824	616
43	CE	144	Sri Lanka	2	24	81	4	14,846	18,322	18,121	1.1	1.68	1.33	1.27	52.29	763	65,610
44	SY	760	Syria	1	13	88	15	11,338	14,045	14,262	1.5	3.46	3.55	3.58	119.51	1,080	185,680
45	TA	762	Tajikistan	1	4	90	21	5,248	5,382	5,933	9.3	0.89	0.84	0.51	189.14	1,346	143,100
46	TH	764	Thailand	1	72	90	11	48,656	57,324	57,586	0.5	1.83	1.32	1.27	84.49	796	514,000
47	TU	792	Turkey	1	73	90	7	56,473	61,301	60,802	0.8	2.50	2.13	2.05	103.33	840	779,450
48	TM	795	Turkmenistan	1	4	90	21	3,622	3,715	4,010	7.4	0.89	0.84	0.51	349.32	929	488,100
49	UA	804	Ukraine	2	25	90	21	51,839	53,165	51,465	3.3	0.89	0.84	0.51	155.40	2,127	603,700
50	TC	784	United Arab Emirate:	1	7	85	16	1,622	2,062	1,861	10.8	3.43	3.26	2.33	103.61	295	75,150
51	UZ	860	Uzbekistan	1	11	90	21	20,322	20,842	22,349	6.7	0.89	0.84	0.51	201.67	1,895	447,400
52	VM	704	Viet Nam	1	39	89	3	69,125	71,215	72,342	1.6	2.18	2.15	2.03	91.93	1,826	329,565
53	YE	887	Yemen	1	2	81	4	9,732.00	15,351	13,873	10.7	3.43	3.60	3.47	488.64	7,676	477,530
					sum	**avg**		**sum**	**sum**	**sum**	**avg**	**avg**	**avg**	**avg**	**avg**	**avg**	**sum**
					4328	**88**		**2,237,563**	**3,641,819**	**3,627,241**	**3.09**	**2.25**	**2.11**	**1.83**	**168.48**	**1,932**	**49,827,394**

Notes By Country:

Former USSR States: Using UN annual rates of change for USSR (former); UN population estimates for 1994 unavailable.

3 AZ Estimate for Azerbaijan adjusted to match UN estimate because of apparent error in source data

11 CY 1992 population figures derived from 1971 provincial population figures.

21 JO Includes West Bank which has CNT-ABB 'WE'. The West Bank population (1,247) was calculated as the difference between the sum of the estimated population for Jordan's provinces and the UN estimate for Jordan

25 KU Using United Nations national estimate

34 BM Chin and Arakan states have only 1st level pop figures though there are 2nd level boundaries. In such cases 2nd level data was unavailable. There are six townships in Shan state for which no population figures were found. A population of -9999 was used to note this.

39 QA Using United Nations national estimate

POPULATION INFORMATION SOURCES

ASIA

1) Munro, D., 1990,. Cambridge World Gazetteer: A Geographical Dictionary, New York, Cambridge University Press.

2) Department of the Army, Cambodia: a Country Study. U.S. Government, Secretary of the Army: Washington D.C., 1990.

3) Viet Nam Statistical Office 1992 estimates based on 1989 census. Unpublished data set obtained from UN Statistical Office, New York.

5) The Europa World Yearbook ,1994, London, Europa Publications Limited.

6) Statistical Centre of Iran, 1990, A Statistical Reflection of The Islamic Republic of Iran. Plan and Budget Organization: Islamic Republic of Iran:

7) State Institute of Statistics, 1991, Statistical Yearbook of Turkey. Prime Ministry Republic of Turkey.

8) Federal Bureau of Statistics: Government of Pakistan, Pakistan Statistical Yearbook, 1989, Manager of Publications, Karachi

9) Demosphere International, India database.

10) The Socialist Republic of the Union of Burma, 1987, Ministry of Home and Religious Affairs, Immigration and Manpower Department, 1983 Population Census

11) National Statistical Office, Office of the Prime Minister, 1990 Population and Housing Census, Thailand., Pho. So. 2533, Bangkok, 1990

12) J.B. Metzler & C.E. Poeschel Verlag, 1989, Landerbericht: Afghanistan 1989, Wiesbaden, Federal Statistical Office.

13) J.B. Metzler & C.E. Poeschel Verlag, 1988, Landerbericht: Irak 1988, Wiesbaden, Federal Statistical Office.

14) J.B. Metzler & C.E. Poeschel Verlag, 1992, Landerbericht: Jordanien 1992, Wiesbaden, Federal Statistical Office.

15) J.B. Metzler & C.E. Poeschel Verlag, 1990, Landerbericht: Syrien 1990, Wiesbaden, Federal Statistical Office.

16) J.B. Metzler & C.E. Poeschel Verlag, 1990, Landerbericht: Vereinegte Arabische Emirate, Wiesbaden, Federal Statistical Office.

17) Statistical Yearbook of Nepal, 1993, Central Bureau of Statistics, Kathmandu, Nepal

18) The Population of North Korea (Population figures for 1987), Eberstatdt, Nichola and Banister, J., 1992, Institute of East Asian Studies: Center for Korean Studies, Berkeley.

19) Korea Statistical Yearbook
Population figures from the Population and Housing Census of 1990
The Korean Statistical Association, Korea, 1992

20) China County and Population Density Digital Map, 1991,
National Geographic Society, Washington D.C., 20036

21) Baseline Demographics for the Republics of the Former Soviet Union, 1992,
New World Demographics.

22) 1990 Philippine Statistical Yearbook, National Statistical Coordination Board
Republic of the Philippines.

23) J.B. Metzler & C.E. Poeschel Verlag, 1990, Landerbericht: Japan 1990, Wiesbaden,
Federal Statistical Office

24) Poston, D. & Yaukey, D., 1992, The Population of Modern China, New York, Plenum Press
(includes figures from 1990 Population Census of China)

25) Indonesia -- Census of Indonesia, 1991, Unpublished data obtained from the UN Statistical Division, New York

26) Preliminary Results of the General Census of Population, Housing
and Establishments, Sultanate of Oman, Development Council, 1993

27) Central Intelligence Agency, 1993, Atlas of the Middle East, Washington D.C., Government Printing Office.

BOUNDARY INFORMATION SOURCES

ASIA

For information on how to obtain any of the data, please contact the address listed under "Contact"

Western Asia

Boundary data in latitude & longitude were taken from ArcWorld's administrative units coverage. Names (where required) and population figures were added to these. 1st subnational level.
1:3 million nominal map scale.
Contact:
ESRI (see USA source)

Pakistan

Digitized at NCGIA from 1:2.5 million scale map: Pakistan Administrative Divisions, 1983. Boundaries were then transformed and projected into geographic units and merged with the Western Asia coverage. Latitude/longitude coordinates. 2nd subnational level. 1:2.5 million base map scale.
Contact: NCGIA (see above)

Saudi Arabia

Map of Saudia Arabia, 1990, scale 1:6,000,000 Washington D.C., Central Intelligence Agency. Digitized at NCGIA.
Contact: NCGIA (see above)

India

Data were obtained from Demosphere International in Arc/Info format. Data received in latitude/longitude coordinates. 2nd subnational level (districts). Base map scale unknown.
Contact: (see Mexico source, above).

Nepal

Data were digitized by Geographic Decision Systems Inc. from a map produced by the Nepal Topographical Survey. Data were projected into geographic coordinates and population data were added to this coverage. Nepal coverage was subsequently merged with the India coverage. Latitude/longitude coordinates. 2nd subnational level (districts). Base map scale 1:2 million.
Contact:
Geographic Decision Systems International
2800 Woodlawn Drive #244
Honolulu, HI 96822-1843
Tel: 808-539-3773, Fax: 808-539-3775

Vietnam & Indonesia

Data were produced by the National Statistical Offices of Vietnam and Indonesia, respectively, in collaboration with the UN Statistical Division's Software development project. Data were obtained in POPMAP format, converted to Arc/Info format and rubber sheeted into geographic coordinates. Coverages were subsequently merged into the South-East Asia coverage. Original projection unknown, transformed into latitude/longitude coordinates. 2nd subnational level. Base map scale unknown.
Contact: (see Cuba source.)

Remaining South East Asia

Data produced by Birbeck College, London, based on World Boundary Data Base II as part of an Indian database in latitude/longitude coordinates. 1st or 2nd subnational level depending on country. 1:3 million nominal map scale.
Contact:
Geography Department, Birbeck College
7/15 Gresse Street, London W1P 1PA, England

China

Data were obtained from the National Geographic Society which used the data for a map accompanying the NGS magazine. The map projection was determined to be Albers Conical Equal

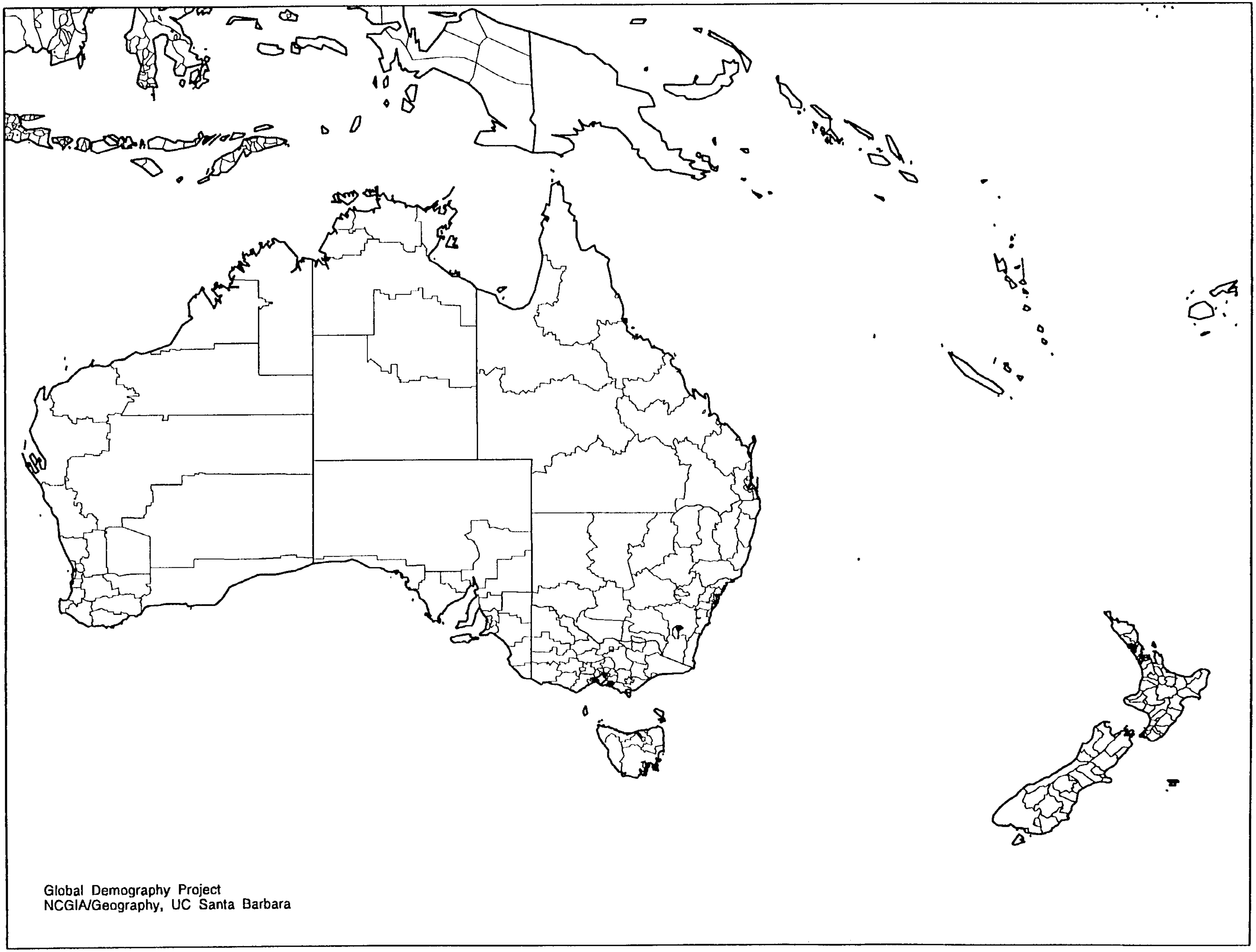
Area Projection. Coordinates were transformed and projected into geographic (lat/lon) projection. 2nd subnational level (counties). Base map scale unknown.
Contact:

National Geographic Society
1145 Seventeenth Street N.W.
Washington, D.C., 20036 USA

Mongolia & North East Asia

Data compiled by GRID Tsukuba, Japan, presumably based on World Boundary Data Base II. Data received in latitude/longitude coordinates. 1st or 2nd subnational level depending on country. 1:3 million nominal map scale.
Contact:

UNEP GRID Tsukuba
Center for Global Environmental Research
National Institute for Environmental Studies
Japan Environment Agency
16-2, Onogawa, Tsukuba
Ibaraki 305, Japan


Global Demography Project
NCGIA/Geography, UC Santa Barbara

Oceania Medium Resolution Population Database — Data Summary Sheet

Num	Abb	UN Code	Country	Level used	Num of units	Ref year	Source	ntry population (in 000) in ref yr (source)	in 1994 (proj.)	in 1994 (UN est)	Abs % diff.	Annual rates of change 80-85	85-90	90-95	Mean resolution in km	Mean pop per unit (in 000)	Area in sqkm
1	AQ	16	American Samoa	0	1	94	3	53	53	53	0.00	3.83	3.70	2.90	14.04	53	197
2	AS	36	Australia	2	187	91	1	17,075	17,827	18,090	1.45	1.40	4.62	1.41	202.69	95	7,682,300
3	CW	184	Cook Islands	0	1	94	3	17	17	17	0.00	-0.36	-0.35	-0.13	15.26	17	233
4	FJ	242	Fiji	0	1	94	3	755	755	755	0.00	1.97	0.77	0.96	135.39	755	18,330
5	FP	258	French Polynesia	0	1	94	3	217	217	217	0.00	2.82	2.56	2.28	62.77	217	3,940
6	GQ	316	Guam	0	1	94	3	143	143	143	0.00	2.46	2.08	1.70	21.21	143	450
7	KR	296	Kiribati	0	1	94	3	77	77	77	0.00	2.06	1.96	2.21	26.15	77	684
8	MH	584	Marshall Islands	0	1	94	3	53	53	53	0.00	3.97	3.64	3.61	13.45	53	181
9	FM	583	Micronesia, Fed. States	0	1	94	3	118	118	118	0.00	2.70	3.39	3.46	26.50	118	702
10	NR	520	Nauru	0	1	94	3	10	10	10	0.00	2.17	2.57	2.21	4.58	10	21
11	NC	540	New Caledonia	0	1	94	3	178	178	178	0.00	1.96	1.62	1.45	138.22	178	19,105
12	NZ	554	New Zealand	2	74	91	2	3,433	3,528	3,520	0.23	0.84	0.87	0.92	59.86	48	265,150
13	NE	570	Niue	0	1	94	3	2	2	2	0.00	-4.68	-5.20	-4.58	16.09	2	259
14	CQ	580	Northern Mariana Islanc	0	1	94	3	47	47	47	0.00	2.89	16.06	2.21	21.70	47	471
15	PW	585	Palau	0	1	94	3	16	16	16	0.00	2.20	1.91	2.21	19.10	16	365
16	PP	598	Papua New Guinea	0	1	94	3	4,246	4,246	4,246	0.00	2.29	2.26	2.29	680.32	4,246	462,840
17	WS	882	Samoa	0	1	94	3	159	159	159	0.00	0.23	0.10	0.15	53.29	159	2,840
18	BP	92	Solomon Islands	0	1	94	3	366	366	366	0.00	3.50	3.42	3.32	172.60	366	29,790
19	TG	776	Tonga	0	1	94	3	98	98	98	0.00	0.49	0.49	0.55	26.44	98	699
20	TV	798	Tuvalu	0	1	94	3	13	13	13	0.00	4.06	4.05	2.90	5.00	13	25
21	NH	548	Vanuatu	0	1	94	3	165	165	165	0.00	2.70	2.45	2.45	121.51	165	14,765
22	WK	872	Wake Island	0	1	94	3	4	4	4	0.00	4.07	3.61	3.60	3.46	4	12
23	WF	876	Wallis and Futuna Islan	0	1	94	3	14	14	14	0.00	2.90	1.29	1.00	15.97	14	255
					sum	**avg**		**sum**	**sum**	**sum**	**avg**	**avg**	**avg**	**avg**	**avg**	**avg**	**sum**
					282	**94**		**27,259**	**28,106**	**28,361**	**0.07**	**2.02**	**2.52**	**1.70**	**80.68**	**300**	**8,503,614**

POPULATION INFORMATION SOURCES

OCEANIA

1) The Europa World Yearbook, 1994, London, Europa Publications Limited.

2) 1991 Census of Population and Housing--Basic Community Profile
Australian Bureau of Statistics, Belconnen ACT

3) 1991 Census of Population and Dwellings
Statistics New Zealand, Wellington.

BOUNDARY INFORMATION SOURCES

OCEANIA

For information on how to obtain any of the data, please contact the address listed under "Contact"

<u>Australia</u>

Data for census sub-divisions were purchased from the Australia Bureau of Statistics and were delivered in Arc/Info format. Data received in latitude/longitude coordinates. 2nd subnational level. Base map scale unknown.
Contact:
Australian Bureau of Statistics
P.O. Box 10
Belconnen, ACT 2616
Tel: 06 252 6972, Fax: 06 253 1404

<u>New Zealand</u>

Digital boundaries for Territorial Local Authorities in Arc/Info format purchased from Statistics New Zealand. Data received in latitude/longitude coordinates. 2nd subnational level. Map scale unknown.
Contact:
Statistics New Zealand
P.O. Box 2922
Wellington, New Zealand
Tel: 64 4 495 4861, Fax: 64 4 495 4617

<u>Remaining Oceania</u>

Boundary data were taken from Arc/World's administrative units coverage. Names (where required) and population figures were added to these. Data received in latitude/longitude coordinates. National level for South Pacific Islands. 1:3 million nominal map scale. Contact: (see USA, above).

SMOOTH MASS PRESERVING AREAL DATA REDISTRIBUTION ON A SPHERE

W. Tobler
Research Professor
NCGIA & Geography Department
University of California
Santa Barbara CA 93106-4060
tobler@geog.ucsb.edu

An objective of the Global Demography Project is to reallocate the number of people recorded within countries to a grid of latitude and longitude quadrilaterals. This is because a reassignment of the populations from the political or census units to a set of spherical cells provides a useful format for many modeling tasks. The data are then also more easily matched to geophysical and earth satellite observations.

In this procedure country or census boundaries are first recorded as polygons defined by latitude & longitude coordinates. The population figures for these units are then assembled. Next a lattice of uniform cells is positioned over the political units, and the population evenly spread over the cells of each unit. This yields what can be thought of as a piecewise continuous surface of population, with discontinuous jumps at the political unit boundaries. To provide more realism a redistribution within each unit is now undertaken by a reallocation designed to give an approximation to a smooth continuous surface of proper content. It considers the populations of adjacent units and the population density transition between units becomes smooth and continuous. This type of data redistribution was introduced for a planar set of regions in a previous publication (Tobler 1979a) and exists in a computer program available from the AAG Microcomputer Specialty Group. The procedure is sometimes known as areal interpolation since the data exist only for areas and are not given at point locations. The present note extends this procedure to a spherical domain. Other geographic applications in which spherical versions of similar equations appear are briefly cited.

The reallocation is required to result in a smooth discretized surface whose volummetric content is identical to the initially given value when integrated over the comparable irregularly shaped region. In the discrete case the integration is of course a summation over the nodes of the grid or lattice. The steps in the procedure are to 1) for each of the separate polygons assign and distribute the entire regional population to the cells of a grid covering that polygon; then 2) reallocate the population smoothly to these same cells in each region, taking into account the population in the neighboring regions, and then 3) test whether the population totals within each region are still correct, and if not, to apply a correction factor. The smoothing uses a converging iterative algorithm that requires many passes over the lattice. The only part of this procedure that requires modification for the spherical case is the smoothing operation in step two, and this is what is covered here.

If one were to apply the procedure to all of South America, using the populations of the countries (defined by polygons bounded by latitude and longitude points) as the data to be redistributed, then it might be thought that an adjustment would be needed for the first step (above) in the procedure. Chile, for example, is a long north-south country and the spherical quadrilaterals covering the country are of unequal area, shrinking in size toward the South Pole. The assignment of the population to the cells could then be adjusted for this latitudinal variation of the latitude/longitude quadrilateral sizes. This does no harm but is not really necessary. The iterative procedure used for the spherical smoothing adjustment overrides the initial pattern. All of the population could actually be assigned to just one cell in the lattice within each country and the final result would be the same, although the iterations might take a bit longer. An exception should be noted here. If the objective is simply to spread the data evenly over the cells of a region representing a polygon, without any subsequent smoothing, then the latitudinal variation in the cell sizes needs to be taken into account. Otherwise the data are not spread evenly, increasing in density as the cells become smaller. To simplify things this adjustment is recommended in all cases. The smoothing step itself does need modification to incorporate the changing spherical cell sizes.

A smooth function, intuitively, is one that has few local oscillations. Stated in another way, neighboring locations should have similar values. Such autocorrelation seems to exist for many human distributions. Alternately one could say that there is only a small rate of change in all directions. Mathematically this requires that the functions' partial derivatives be small. Thus it is natural to minimize the sum of the squares of the partial derivatives, i.e., minimize:

$$\iint \left(\frac{\partial U^2}{\partial x} + \frac{\partial U^2}{\partial y} \right) dx\, dy \quad .$$

Here the integral is to be taken over the entire region of interest, and U(x,y) is the function in question. The solution to this least squares problem is another equation, known as the Laplace equation (Kantorovich and Krylov 1958, p 246 et seq.):

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} = 0 \ .$$

Solving this new partial differential equation, with appropriate boundary conditions, gives the required smooth function U(x,y). To solve the equation on a digital computer it is usual to discretize the equation and to work with a finite lattice. The finite difference version to Laplace's equation is obtained by using the well known approximations

$$\frac{\partial^2 U}{\partial x^2} \approx U_{i,j-1} - 2U_{i,j} + U_{i,j+1} \ , \quad \frac{\partial^2 U}{\partial y^2} \approx U_{i-1,j} - 2U_{i,j} + U_{i+1,j} \ .$$

Here we use i for a row subscript and j for a column subscript with the value at each node of the mesh labeled U and subscripted by its row and column position. Adding these two equations and setting the result equal to zero, yields the condition that the value at any mesh point must be equal to the average of its neighbors. This is another way of defining what is known technically as a 'harmonic' function. Expressed in symbols, using four grid neighbors, this gives the equation:

$$U_{i,j} = \tfrac{1}{4}(\ U_{i-1,j} + U_{i+1,j} + U_{i,j-1} + U_{i,j-1}\ )$$

An even stronger requirement is that some derivative be the average of its neighbors. This results in a higher order of smoothness. The usual approach is to use the Laplacian of Laplace's equation. After some tedious but simple algebra and taking the grid spacing $\Delta x = \Delta y \equiv 1$, the required condition becomes

$$U_{i,j} = [8(U_{i+1,j} + U_{i-1,j} + U_{i,j-1} + U_{i,j+1}) - 2(U_{i+1,j+1} + U_{i-1,j+1} + U_{i+1,j-1} + U_{i-1,j-1}) - (U_{i+2,j} + U_{i,j+2} + U_{i-2,j} + U_{i,j-2})] \ / \ 20 \ .$$

This is a finite difference approximation to the biharmonic equation

$$\frac{\partial^4 U}{\partial x^4} + 2\frac{\partial^4 U}{\partial x^2 \partial y^2} + \frac{\partial^4 U}{\partial y^4} = 0$$

which solves a least squares problem minimizing the surface curvature. These are the two alternate smoothing operations proposed and previously described for the mass preserving interpolation problem on a plane. Boundary conditions are also needed for a solution.

The Laplace equation on the surface of a unit sphere is

$$\frac{1}{\cos\phi}\frac{\partial}{\partial\phi}\left(\cos\phi\frac{\partial U_{\phi,\lambda}}{\partial\phi}\right) + \frac{1}{\cos^2\phi}\frac{\partial^2 U_{\phi,\lambda}}{\partial\lambda^2} = 0$$

$$\phi = \textit{latitude} \quad , \quad \lambda = \textit{longitude}$$

and the spherical surface biharmonic equation is

$$\frac{1}{\cos^2\phi}\frac{\partial}{\partial\phi}\cos\phi\frac{\partial^2}{\partial\phi^2}\cos\phi\frac{\partial U}{\partial\phi} + \frac{1}{\cos^3\phi}\left(\frac{\partial^3}{\partial\phi\partial^2\lambda}\cos\phi\frac{\partial U}{\partial\phi} + \frac{\partial}{\partial\phi}\cos\phi\frac{\partial^3 U}{\partial\phi\partial^2\lambda}\right) + \frac{1}{\cos^4\phi}\frac{\partial^4 U}{\partial\lambda^4}$$

To get the spherical versions in discrete form, begin with the harmonic property and modify the finite difference approximation to the Laplace equation so that it can be applied on the surface of a sphere. That is, use the fact that the value at any location is the average of its neighbors. For this purpose a weighted average of the neighbors must be used because the neighbors are at different spherical distances in the latitude/longitude graticule.

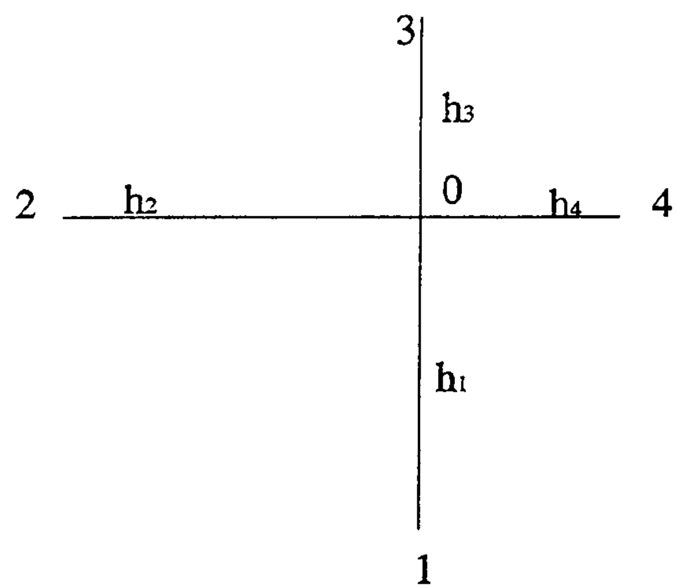
The general centered difference formula for the Laplace equation when the central point at u, is related to neighbors at different distances is (Rektorys 1969, p. 1114)

$$\frac{2}{h_1+h_3}\left(\frac{u_1-u_0}{h_1}+\frac{u_3-u_o}{h_3}\right)+\frac{2}{h_2+h_4}\left(\frac{u_2-u_0}{h_2}+\frac{u_4-u_0}{h_4}\right)=0\ ,$$

and is explained by reference to the diagram on the next page. For the spherical case the distances h, and h3 along the meridian are equal to each other, and h2 and h4 (along a parallel) are equal to each other but these are not equal to the meridional distance. Let h1 = h3 = h and h2 = h4 = h cos(ϕ), from the geometry of the sphere. Then the sought after average value at Uij is obtained from some simple algebra and is, reverting to the row, column notation:

$$U_{i,j}=\frac{1}{2}[U_{i,j-1}+U_{i,j+1}+\cos^2\phi\,(U_{i-1,j}+U_{i+1,j})]\,/\,(1+\cos^2\phi)\ .$$

The divisor is a normalization factor to take into account the unequal weights in the weighted average.



This agrees with the result by Swarztrauber (1974), and approximates that of Gates & Riegel (1963), if we consider that the values apply to small quadrilaterals not near either pole. It assumes that the cosine values vary by only a small amount over the small increment. The cosine value should be taken at the center of the cell and not at the edge. These equations further assume a sphere of unit radius and equal increments in latitude and longitude. If the spacing is not equal or the cells are of large size then the more general equations must be substituted. The biharmonic case takes a bit more algebra but uses the same logic. A similar, slightly more complicated, modification is required to adapt the equations to an ellipsoid.

In the computer program one needs to know the latitude of the southernmost (or northernmost) mesh line and the latitudinal increment between these lines. A table of squared cosines and divisors is precomputed so that these do not need to be recalculated on every iteration. If the cell centers are to contain the information then these cosine values should be shifted by one half the increment. It is also necessary to know whether the domain extends over 360 degrees of longitude. If it does then the computation must recognize the periodic nature of the east-west boundary. Otherwise a Dirichlet or Neumann condition must be specified at all edges. In the context of the current application the grid will not reach the poles since there will be no data of relevance there.

The smooth pycnophylactic; reallocation for areal data can also be used as one step in the more general conversion of values from one set of polygons to another, e.g., converting information from political units to ecological zones, or vice versa. The modification described here allows this to be done when the two sets of polygons are defined by latitude/longitude coordinates. Several additional uses of these equations can be noted. In Tobler & Kennedy (1985) a procedure was described in which the Laplace equation (or the biharmonic equation) is used to solve the interpolation problem for randomly scattered points on a plane. Results given here allow this interpolation technique to be applied to data irregularly arranged on a sphere. Also see Tobler (1979b and 1994) for related uses of smooth interpolation from values given at point locations. In Dorigo and Tobler (1983) an algebraic model of geographic movement is developed which allows estimation of a potential function and associated vector field from empirical data by solving Poisson's equation. The finite difference approximation to the Poisson equation is identical to that used for the Laplace equation aside from the term- defining the source/sink field. Using $f(\phi,\lambda)$ as this forcing function one obtains

$$U_{i,j}=[U_{i,j-1}+U_{i,j+1}+\cos^2(\phi)(U_{i+1,j}+U_{i-1,j})+\cos^2(\phi)f(\phi,\lambda)] \;/\; [2(1+\cos^2(\phi))].$$

The push-pull model can therefore be applied to movement data on the surface of the earth, assumed spherical, with observations in a small mesh of uniform latitude/longitude increments.

References:

AAG Microcomputer Specialty Group, 1994, Spatial Sciences Center, Indiana University of Pennsylvania, 212 Whitmyre Hall, Indiana, PA 15705-1087.

Dorigo, G., and W. Tobler, 1983, "Push-Pull Migration Laws", *Annals, Assoc. Am. Geographers*, 73(l):1-17.

Gates, W., & C. Riegel, 1963, "Comparative numerical integration of simple atmospheric models on a spherical grid", *Tellus*, XV:410.

Kantorovich, L., and V. Krylov, 1958, *Approximate Methods of Higher Analysis*, Nordhoff Interscience, The Hague.

Rektorys, K, ed., 1969, *Survey of Applicable Mathematics*, MIT Press, Cambridge.

Swarztrauber, P., 1974, "The Direct Solution of the Discrete Poisson Equation on the Surface of A Sphere", *J. Computational Physics*, 15:46-54.

Tobler, W., 1979a, "Smooth Pycnophylatic Interpolation for Geographic Regions", *J Am. Stat. Assn.*, 74(367):519-536.

Tobler, W., 1979b, "Lattice Tuning", *Geographic Analysis*, 11(l):36-44

Tobler, W., and S. Kennedy, 1985, "Smooth Multidimensional Interpolation", *Geographical Analysis*, 17(3):251-257.

Tobler, W., 1994, "Bidimensional Regression", *Geographical Analysis*, 26:186-212.