Global Population Distribution Analysis

In partial fulfillment of Advanced GIS 6513
Department of Earth and Environmental Science
University of Texas at San Antonio

Amber Fisher and Sergio Martinez

ABSTRACT

Population distribution has been long studied. There are no clear answers as to what the major influential factors that drive population density are. Using ArcGIS 9.1, this study focuses on global population distribution and it's relationship to selected biological and environmental factors. Using geoprocessing techniques such as reclassify, weighted overlay and modelbuilder, models where created for both environmental factors and biological factors to determine which had greater influence in population density on a global scale. Statistics were then run on the fifteen countries with the highest population density and lowest population density using both the environmental weighted raster and biological weighted raster. Results showed that the chosen biological factors had a greater influence on population density when compared to the environmental factors used for this study.

INTRODUCTION

As of late 2006, the world population reached 6.5 billion. In line with population projections, this figure continues to grow at rates that were unprecedented prior to the 20th century, although the rate of increase has almost halved since growth rates reached their peak in 1963. Different regions have different rates of population growth and different population densities. The 20th century saw the biggest increase in the world's population in human history.

It is important to understand how and where the population is distributed throughout the world since populations are not uniformly distributed on Earth's landmass and neither are physical environments. It is also important to understand the effects of different changes in population on the world. It is useful to understand spatially where the population is located presently and where it may move in the future, since the future of population of the world is difficult to predict. By understanding certain patterns and understanding which factors have the heaviest impact on population density and total population, it is possible to predict the future growth or possible decline in global population, and also locate them spatially throughout the world. Then it will be possible to develop policies that can protect the environment, allow for sustainability of the planet, and allow the continued change of global population.

This project focuses on global population density in relation to demographic figures and environmental factors. It will attempt to establish a relationship between different variables and population and to find patterns that exist on a global scale. Looking at the different characteristics of population worldwide, such as birth rates, fertility rates, death rates and population density across the world, this project will analyze if the variables are directly linked to the spatial geography of population, and if it is influenced by environmental factors such as agricultural land cover, total forest extent, hazard frequency and distribution, annual precipitation, and natural renewable water resources.

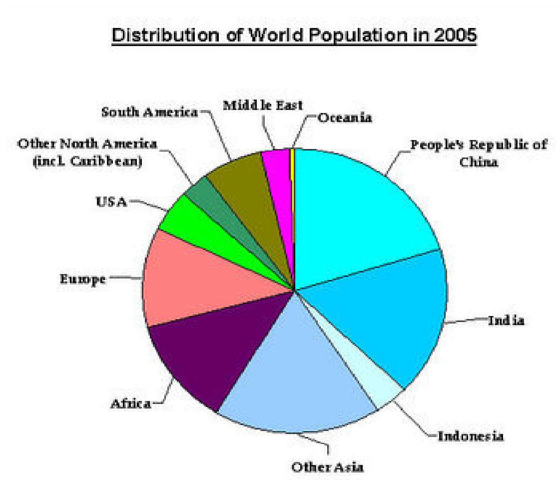
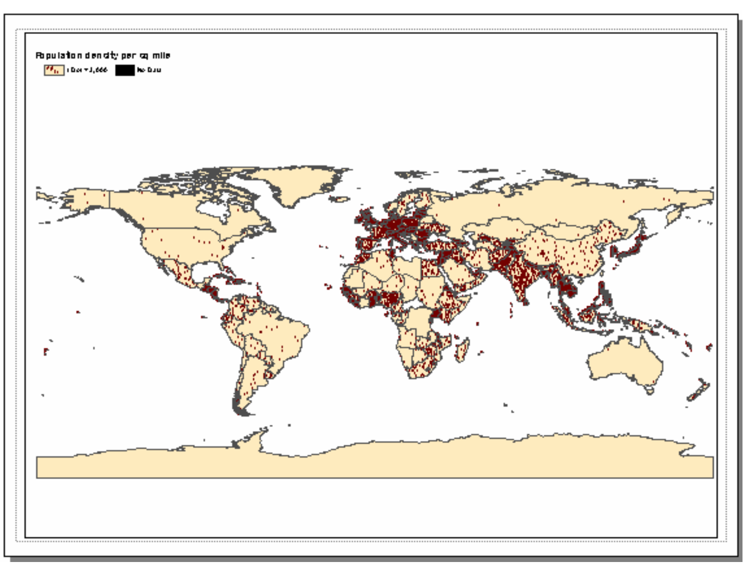


Fig. 1 Population by region, 2005
Source: http://en.wikipedia.org/wiki/Global_Population#Rate_of_increase

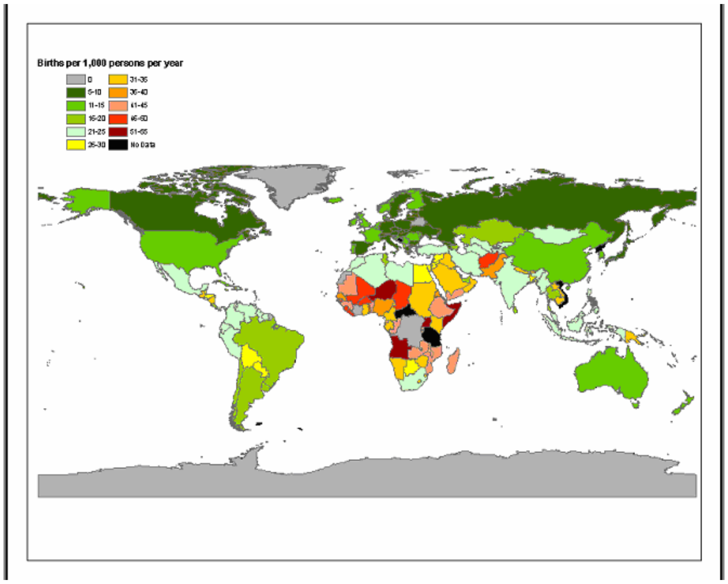
DATA

A country shapefile from the International Potato Center was used for the base map of this study. Statistical data used for a nonspatial table was compiled from Population Reference Bureau World Population Data Sheet, and WDI Online (World Development Indicators Online). All statistical data used was from the year 2003. Data found was for each country as a whole (although some countries did not have any data available), mostly dealing with population characteristics. The statistics used were population density, crude birth rate, crude death rate, fertility rate, and annual population growth rate, shown in the following maps.

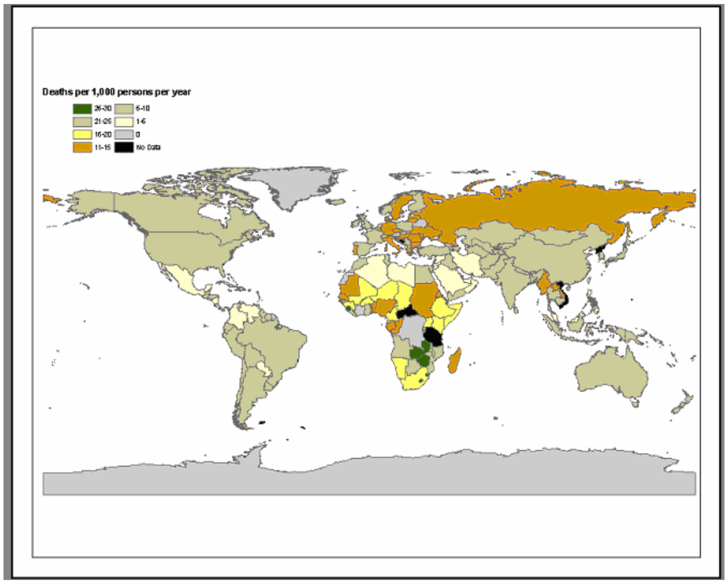
Population Density



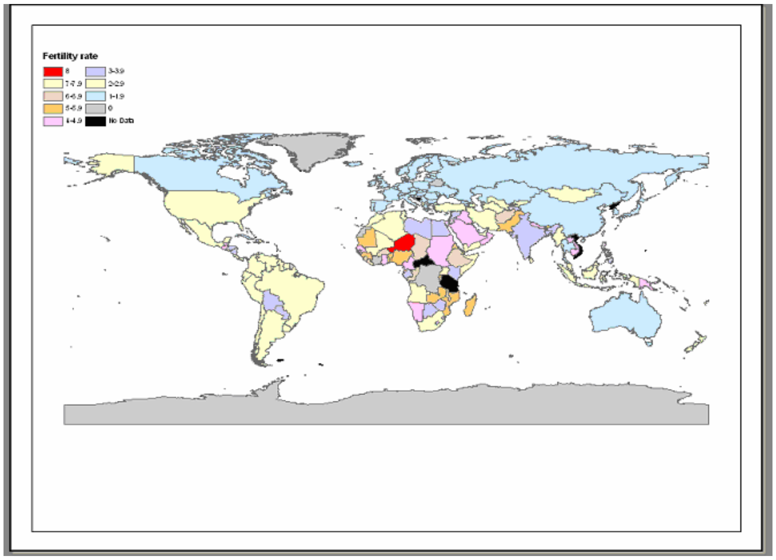
Crude Birth Rate



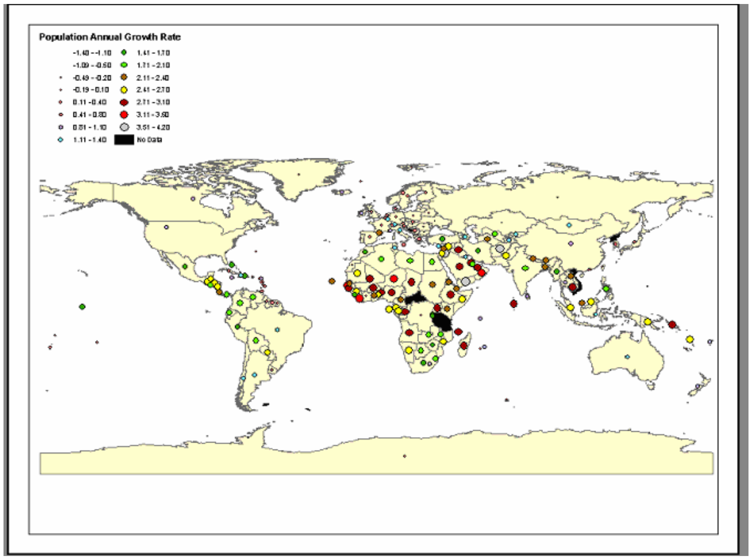
Crude Death Rate



Fertility Rate



Population Annual Growth Rate



Other data used included the following shapefiles:

- ***Natural Disasters- # of total affected and killed people*** is a world coverage national dataset compiled from various sources, including UN agencies, non-governmental organizations, insurance companies, research institutes and press agencies. The time period covered in this dataset is 1900-2002.

Definitions for this dataset are as follows:
Total affected: People that have been injured, affected and left homeless after a disaster are included in this category. Killed: Persons confirmed as dead and persons missing and presumed dead. Disaster: Situation or event, which overwhelms local capacity, necessitating a request to national or international level for external assistance; an unforeseen and often sudden event that causes great damage, destruction and human suffering. Though often caused by nature, disasters can have human origins. Wars and civil disturbances that destroy homelands and displace people are included among the causes of disasters. Other causes can be: building collapse, blizzard, drought, epidemic, earthquake, explosion, fire, flood, hazardous material or transportation incident (such as a chemical spill), hurricane, nuclear incident, tornado, or volcano. Natural disasters include: Droughts, Earthquakes, Extreme Temperatures, Famine, Floods, Insect infestation, Slides, Volcanic eruptions, Wave/surges, Wild fires and Wind storms. Note that grid cells of "No Data" are not necessarily without risk from hazards, but may rather be an artifact of a mask that excludes from.

- ***Mean Annual Precipitation units in millimeters*** is a world coverage geospatial dataset from the United Nations Environment Programme's GEO Data Portal. Mean annual precipitation is the annual average (within 1961-90) of water falling on the country. Data was obtained through a network of global climate stations and the final data set established contains 3578 precipitation means. Interpolation of the precipitation data to a grid-cell map was done using the simplest and most robust technique available; that is, a distance-weighted, nearest-neighbor interpolation. In this technique, a value was computed as an average of 'k' nearest station values ('k' typically equaled 4) and weighted by the inverse distance squared. The data was then converted to Arc/INFO vector format. The final Mean Annual Precipitation data set derived contains a total of 672 polygons, and shows mean annual precipitation (in mm.) for the period 1961-1990.

- ***Total Forest Extent units in thousand hectares*** is a world coverage sub regional dataset with world coverage data from the United Nations Environment Programme's GEO Data Portal. Information has been collated from 229 countries and 2000 territories. This information examines the current status and recent trends for about 40 variables covering the extent, condition, uses and values of forests and other wooded land. The results are presented according to six thematic elements of sustainable forest management. This dataset covers the year 2002.

The definition of forest for this purpose is:
1. Land spanning more than 0.5 hectares with trees higher than 5 meters and a canopy cover of more than 10 percent, or trees able to reach these thresholds in situ. It does not

include land that is predominantly under agricultural or urban land use. Forest is determined both by the presence of trees and the absence of other predominant land uses. The trees should be able to reach a minimum height of 5 meters in situ. Areas under reforestation that have not yet reached but are expected to reach a canopy cover of 10 percent and a tree height of 5 m are included, as are temporarily unstocked areas, resulting from human intervention or natural causes, which are expected to regenerate.

2. Includes areas with bamboo and palms provided that height and canopy cover criteria are met.

3. Includes forest roads, firebreaks and other small open areas; forest in national parks, nature reserves and other protected areas such as those of specific scientific, historical, cultural or spiritual interest.

4. Includes windbreaks, shelterbelts and corridors of trees with an area of more than 0.5 ha and width of more than 20 m.

5. Includes plantations primarily used for forestry or protection purposes, such as rubberwood plantations and cork oak stands.

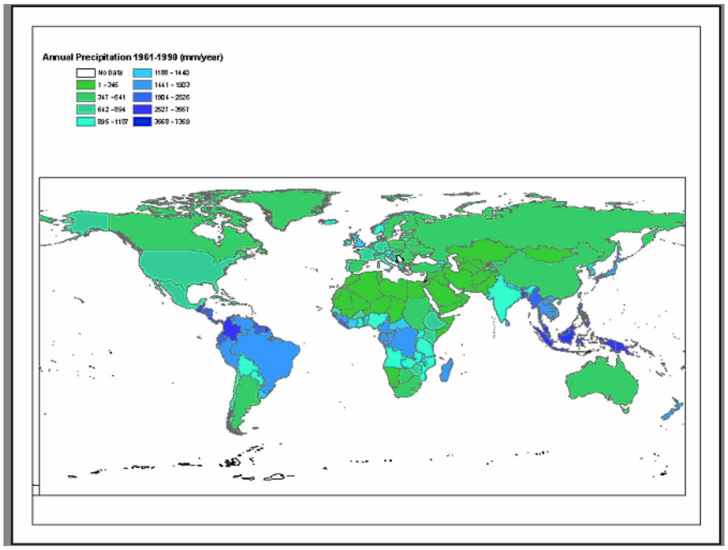
6. Excludes tree stands in agricultural production systems, for example in fruit plantations and agroforestry systems. The term also excludes trees in urban parks and gardens.

- *Agricultural units in thousand hectares* is a world coverage sub regional dataset with world coverage data from the United Nations Environment Programme's GEO Data Portal.

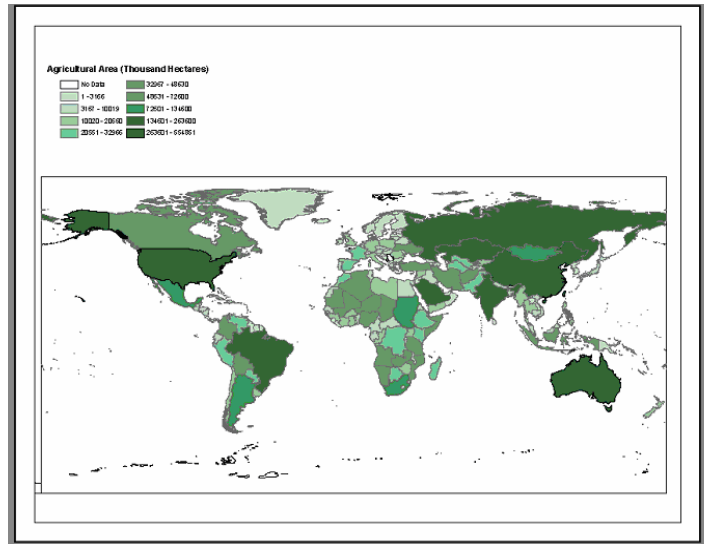
Definition for this dataset is as follows:
Agricultural Area for this purpose is the sum of Arable Land and Permanent Crops, plus Permanent Pastures. Arable Land: land under temporary crops (double-cropped areas are counted only once), temporary meadows for mowing or pasture, land under market and kitchen gardens and land temporarily fallow (less than five years). The abandoned land resulting from shifting cultivation is not included in this category. Data for "Arable land" are not meant to indicate the amount of land that is potentially cultivable. Permanent Crops: land cultivated with crops that occupy the land for long periods and need not be replanted after each harvest, such as cocoa, coffee and rubber; this category includes land under flowering shrubs, fruit trees, nut trees and vines, but excludes land under trees grown for wood or timber. Permanent Pasture: land used permanently (five years or more) for herbaceous forage crops, either cultivated or growing wild (wild prairie or grazing land). This dataset covers the year 2000.

- *Natural Renewable Water Resources per m^3 per person* is world coverage data from the United Nations Environment Programme's GEO Data Portal. Natural renewable water resources are the sum of internal and external renewable water resources. It corresponds to the maximum theoretical amount of water available for a country on an average year on a long reference period. This dataset covers the year 2002.

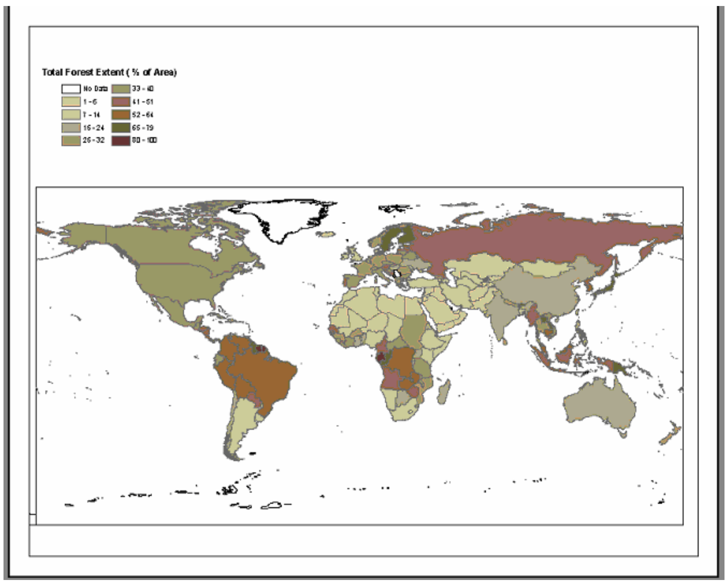
Annual Precipitation



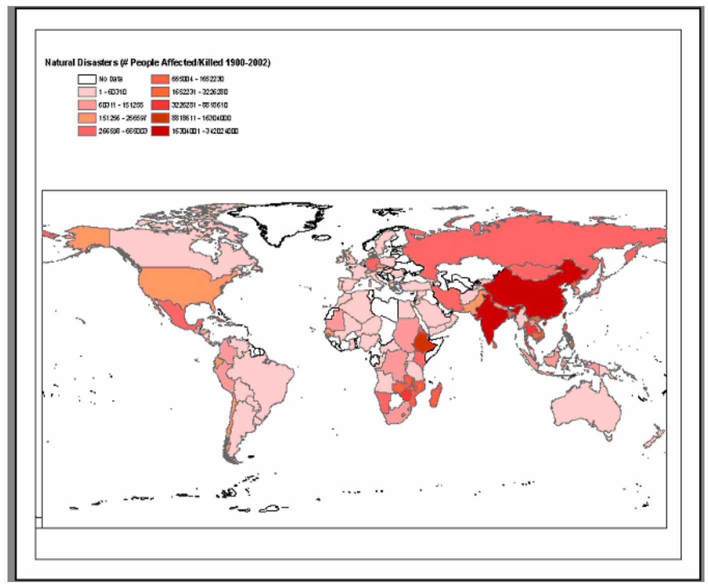
Agricultural Land Cover



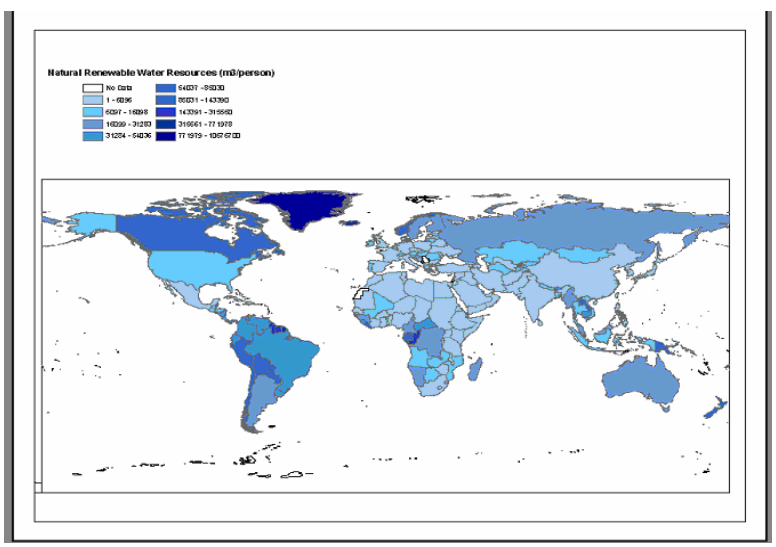
Total Forest Extent



Natural Disasters



Natural Renewable Water Resources



These maps show the data/parameters that are important in analyzing global population. The variations amongst the population and the spatial geography show that patterns and relations do exist between the different categories of data.

METHODS

The first step in analysis was Reclassification, which is a technique for applying a common measurement scale of values to diverse and dissimilar inputs to create an integrated analysis. The next step was Weighted Overlay.

In Weighted Overlay percent values were assigned to each reclassified raster as follows to produce two composite output rasters; environmental and biological:

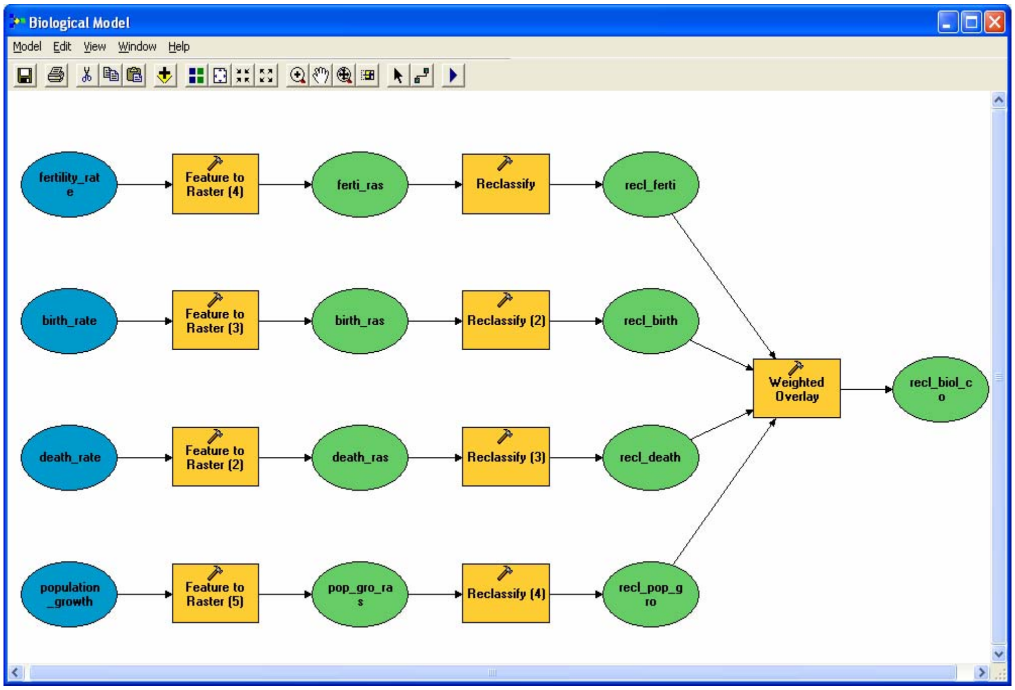
Table 1: Raster Weighted Values

Environmental Composite Raster		Biological Composite Raster	
Annual Precipitation	35%	Crude Birth Rate	25%
Agricultural Land Cover	20%	Crude Death Rate	25%
Total Forest Extent	5%	Fertility Rate	10%
Natural Disasters	10%	Population Annual Growth Rate	40%
Natural Renewable Water Resources	30%		

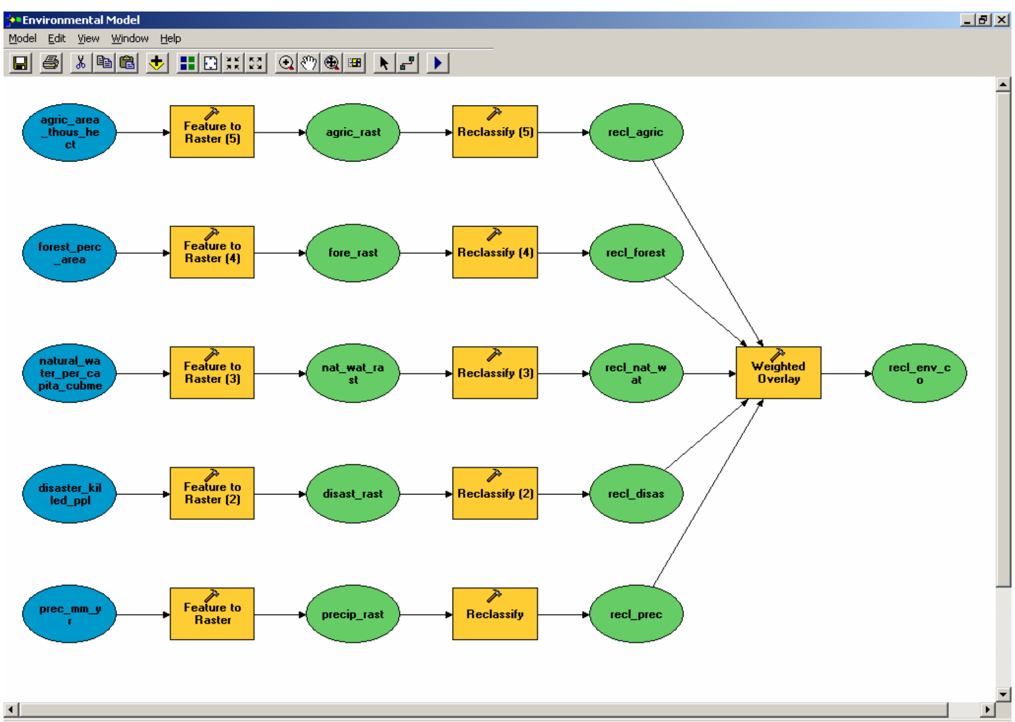
Weighted Overlay only accepts integer rasters as input, such as a raster of land use. Continuous (floating point) rasters must be reclassified as integer before they can be used. The values of the continuous rasters were grouped into ranges. Each range was assigned a single value before it was used in the Weighted Overlay tool. The Reclassify tool allowed for such rasters to be reclassified. Since the factors in the analysis were not equally important, values were prioritized. The evaluation scale was from 1 to 5, a scale value of 5 to the input value of 1 (the most influential factor), a scale value of 4 to the input value of 2 (the second most influential) and a scale value of 3 to the input value of 3 (less influential) and so on. Setting a scale value to 'Restricted' was used to exclude areas of NoData. Then with the correct evaluation scale chosen, the rasters were added to the Weighted Overlay dialog box. The output rasters were weighted by importance (Table 1) and added to produce an output raster. Since the tool was used for a type of suitability modeling (to locate areas were populations would ideally exist), higher values generally indicate that a location is more ideal.

ModelBuilder was used to construct spatial models for this study. The models are represented as process flow diagrams. The flow diagrams created with ModelBuilder were a convenient way to construct and modify the spatial models and implement the geoprocessing tools for workflow. Below are the two models that were constructed with ModelBuilder followed by the final raster output maps for each model.

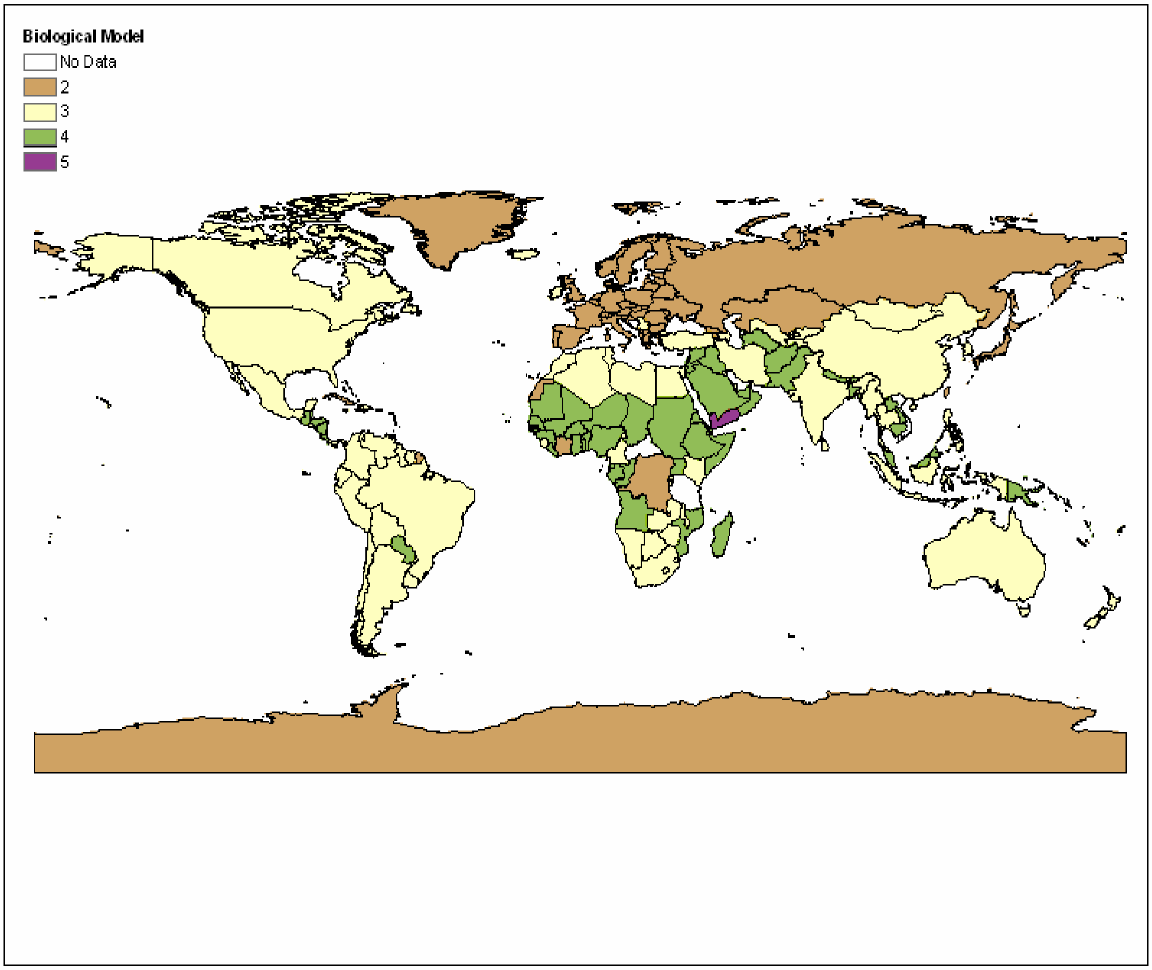
Biological Model



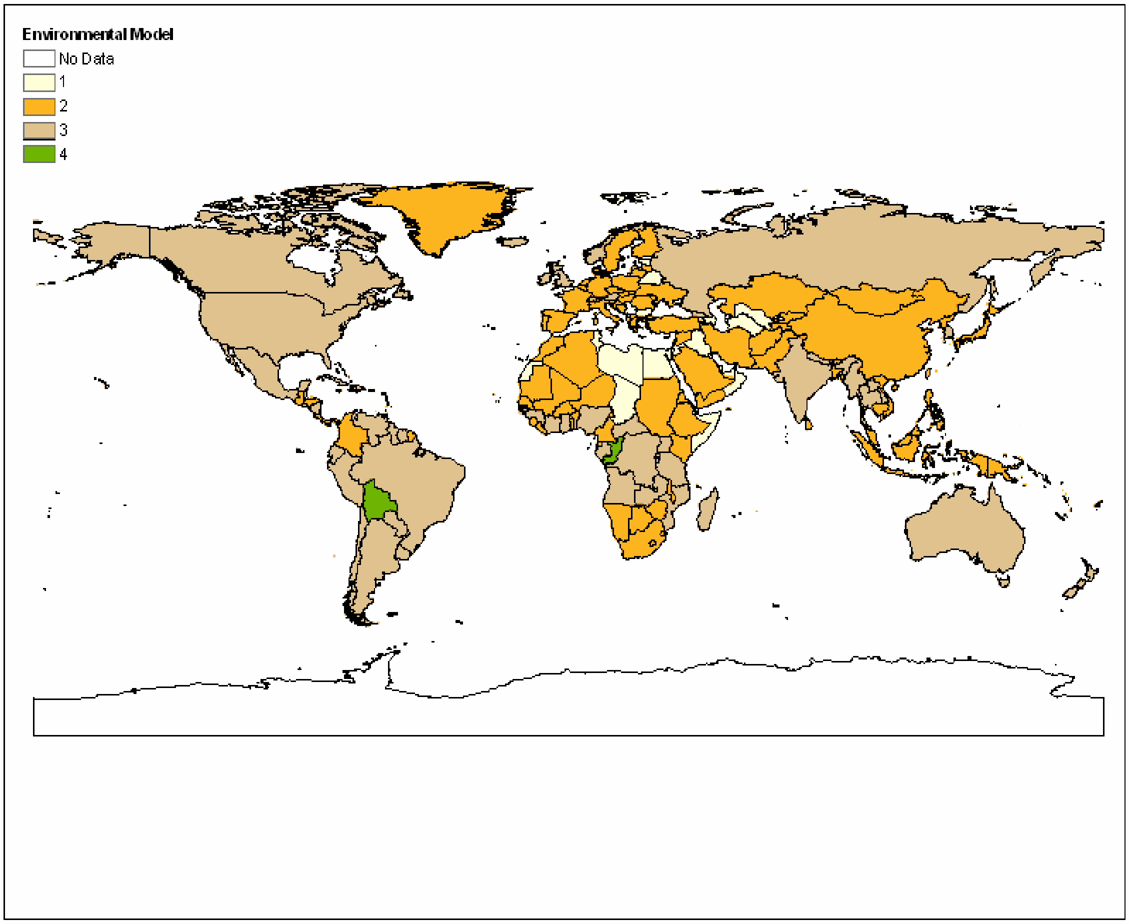
Environmental Model



Biological Composite Output Map



Environmental Composite Output Map



With the final raster composite output maps, population density from the following countries were evaluated to see at which scale values (1-5) these populations exist and at what values do these populations tend to concentrate: Monaco, Singapore, Malta, Bahrain, Bangladesh, Occupied Palestinian Territory, Mauritius, Nauru, San Marino, Republic of Korea, Netherlands, Tuvalu, Lebanon, Marshall Islands, Belgium (15 countries with the highest population density) and Mongolia, Namibia, Australia, Botswana, Canada, Iceland, Mauritania, Suriname, Libya, Gabon, Kazakhstan, Central African Republic, Chad, Bolivia, Russia (15 countries with the lowest population density). This evaluation was performed for both the biological composite output raster and environmental composite output raster.

Averages for the scale values were calculated for the fifteen countries with the highest population density and the fifteen countries with the lowest population for both biological and environmental composite output rasters. It was then determined which model (raster) has a greater influence on population distribution by determining if greater concentrations of population density existed in higher value areas (4 and 5 on evaluation scale) in the environmental raster or if a greater concentration of population density exists in higher value areas of the biological raster.

RESULTS

In the preliminary research it was found that population is very weakly related to land area. For example, Sudan, which is geographically the largest country in Africa, has a smaller population than Nigeria, Egypt, Ethiopia, Democratic Republic of Congo, South Africa or Tanzania and population also decreases rapidly with distance from coastlines and with increase in elevation.

The data regarding the different statistics of a country showed that there was a relationship between the different variables. The data showed that birth rates, death rates, population annual growth rate and fertility rates all correlate. Areas that have high birth rates are likely to have high death rates and high fertility rates, along with high annual population growth rates whereas; areas that have low birth rates are likely to have low death rates, low fertility rates, and low annual population growth rates as well.

Population is more evenly distributed with respect to range of precipitation than any other parameters chosen in this study when parameters were evaluated on an individual basis. The global distribution of population with respect to environmental parameters however, is not strongly localized when compared to the biological parameters chosen for this study when comparing the final raster composite output maps.

In the final raster composite output map for biological parameters, the value '1' on the evaluation scale was not a value determined/displayed in the weighted overlay. This is due to the fact that according to the biological parameters chosen for this study there is no place in the world that would not be and is not completely uninfluenced by these biological factors since population exists in every country

In the final raster composite output map for environmental parameters, the value '5' on the evaluation scale was not a value determined/displayed in the weighted overlay. This is due to the fact that according to the environmental parameters chosen for this study there is no place in the world that is completely influenced by the environmental parameters chosen for the study combined, according to their weighted values.

The average scale values calculated for the fifteen countries with the highest population density was 3.27 for the biological model and 2.20 for the environmental model. The average scale value for the fifteen countries with the lowest population density was 3.00 for the biological model and 2.40 for the environmental model. The low population density countries were used as control to assess if the average scale values for the highest population density countries correlated with the results of the lowest density countries' populations.

These results show populations are more strongly localized by biological factors (3.27) than environmental factors (2.2) on a global scale. This statement can be confirmed since both highly dense populations (3.27 biological / 2.2 environmental) and low-density populations (3.0 biological / 2.5 environmental) are both influenced greater by biological factors than environmental factors. Climate is obviously a very important influence on human settlement but it is presumably less of a constraint on dense settlements than on sparse populations especially when elevation is not being taken into consideration.

CONLUSIONS

The results of the study did not go as expected, as it was assumed that there would be a greater correlation between the distributions of population to the variables of environmental factors. This analysis showed that the density of population and the variables correlated with it are heavy factors in the demography of countries. This project reinforces the idea that different social, biological, economic, political along with environmental factors are more likely to shape the different sets of statistic regarding the population rather than just environmental factors itself, especially the size of the population. Many of the different variables corresponded to the other variables, as a lot of measures of human population and trends are related to one another.

Globally, the growth rate of the human population has been steadily declining (ie. population is growing more slowly than in the recent past), but growth remains high in the Middle East and Sub-Saharan Africa. In some countries there is negative population growth (ie. net decrease in population over time), especially in Central and Eastern Europe (mainly due to low fertility rates) and Southern Africa (due to the high number of HIV-related deaths).

Populations in certain regions will grow; elsewhere, human numbers will stabilize or even decline. Within countries, populations will continue to shift from rural to urban areas, while becoming increasingly older and better educated. Migration between countries will be an increasingly important factor in international relations and the composition of national populations. A global trend towards urbanization also is taking place. The world's urban population is growing by 60 million a year, about three times the increase in the rural population. Population growth rates are also highly dependent upon level of development.

It must be taken into account that any biological process may be affected by multiple environmental factors and visa-versa; there is no one-to-one correspondence between factors and processes (Fig. 2). With this in mind our study found that with no one-to-one correspondence existing between the two. Population density cannot be heavily influenced by only just biological factors or only by environmental factors but a combination of both. It could also be that environmental change is linked in part to populations' change and visa versa.

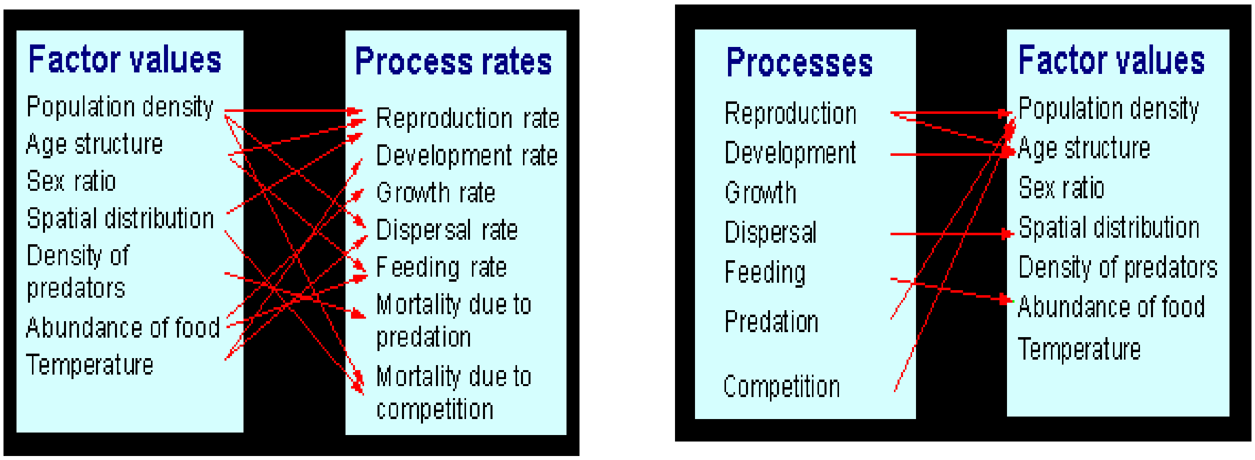


Fig. 2 Correspondence of factor values and processes

The population–environment relationship is about how humans interact with their environment – how they affect and in turn are affected by the environment. A basic diagram for the relationship between humans and environmental conditions is as follows:

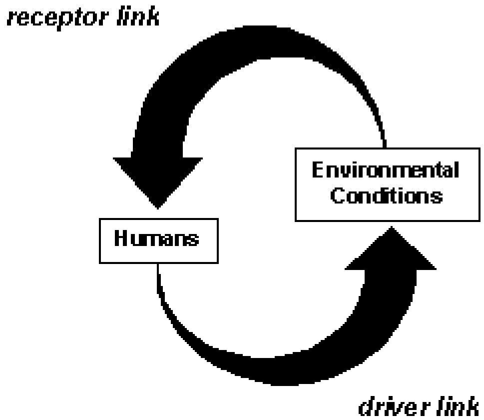


Fig. 3 Human and environmental conditions relationship

In this diagram, the arrows indicate that each side affects the other side, while recognizing the dynamic interplay between population variables and the environment. For example, as population density increases, birth rates often decrease and death rates typically increase. This change in population can affect certain environmental conditions or factors. True can be said for environmental factors on population density. So even though the study concluded that biological parmeters might way somewhat heavier in global population distribution, further evaluation is needed to take into account the very complex relationship of human biological factors and environmental conditions.

IMPROVEMENTS

Improvements for this study could be made in the area of error and uncertainty. County data was gathered at the country level, rather than on a smaller scale like province or region without taking into consideration the international borders. This would provide more accurate information regarding the spatial geographic population if administrative units were smaller.

The dot maps, which showed population density, are also imprecise, as they are randomly placed within the areal unit of the country without regard to where the population may actually lie. The data for all other parameters could also be imprecise, as countries were assigned single values for the data or each parameter used when each country is very dynamic. For example the environmental parameters used could have wide ranges or vary greatly within a single country but one value was assigned to represent the average within that country.

There could also be error in the difference between the dates of the data, as the statistics from the countries were from 2003 while the shapefiles were from the years 2000-2002. If a country's status/statistics changed within that two to three year period, then the map would not show those results.

The reclassification method of the data could also misrepresent the data. The reclassification method could have been somewhat subjective; even though the data was reclassified after thorough research of other modeling studies of global population distribution. This could mean that had the map been reclassified in a different way, the results could have differed significantly. This is also true for the weighted values (%) that were assigned to the reclassified rasters.

FUTURE RESEARCH

Another method that could be used in future studies that might result in a different or maybe similar outcome would be regression analysis. Regression analysis is a multivariate analysis that establishes relationships between numerous input variables and presents the relationships in a succinct manner, usually as a number or a series of numbers.

A regression analysis has two parts:

- The dependent variable, which is the phenomenon whose level or presence you are trying to predict or explain for each location in a study site.
- The independent variables, which are the known attributes of the locations that influence the level or presence of the dependent variable.

Many relevant independent variables can be used in this analysis. Relevant independent variables are any attribute variables that have a bearing on the quantity or probability of the dependent variable (in this example, the amount of biomass). Discrete data types, except those describing the presence or absence of a variable, such as landuse, are not proper sources of data for regression analysis. So this type of method would be useful in evaluating parameters such as birth rate, death rate, fertility rate, net migration, population density, urban population, rural population, ect.

This project could also be continued by looking at many other environmental effects on global population such as the possible threat of the rise in sea level in the future from global warming, which could displace population and cause forced migration, along with many other environmental factors including elevation. This study could also be taken further by looking at other demographic parameters that are available such as infant mortality, GDP per capita, population annual growth rate, and net migration to see how those factors contribute to the distribution of population at the global level. It would also be interesting to compare the differences of global population between different years, and to note the changes or make predictions for the future demographics of global population.

ACKNOWLEDGEMENTS

We would like to thank Dr. Xie for the opportunity to work on a study of our choice in order to extend our knowledge of GIS and also his advice in study methods. We would also like to thank Xianwei for his time in making the computer lab accessible throughout the semester so we could do our research.

REFERENCES

Berry, L. and Johnson, D. (1986) Geographical approaches to environmental change: assessing human impacts on global resources. In *Natural resources and people: conceptual issues in interdisciplinary research,* pp. 67-105. Boulder, CO: Westview Press.

Clark, W.C. The human ecology of global change. *International Social Science Journal*, **41** (1989), pp. 315-345.

Kapitza, S.P., Mathematical model of world population growth, *Math. Model.* **4** (1992), pp. 65–79.

Podlazov, 2002 A.V. Podlazov, Theoretical demography. Models of population growth and global demographic transition, *Advances in Synergetics: The Outlook for the Third Millennium*, Nauka, Moscow (2002) pp. 324–345.

Population Reference Bureau. URL accessed on January 22, 2007. http://www.prb.org/datafind/datafinder7.htm.

PRB, 2004 PRB, World Population Data Sheet, Population Reference Bureau, Washington, DC (2004).

Rogerson, Peter A., The Future of Global Population Modeling, *Futures.* **29** (1997), pp. 381-392.

Severtsov, A.S., Human population dynamics from the viewpoint of population ecology, *Bull. Moscow Soc. Investigators Nat. Biol. Dept.* **97** (1992), pp. 3–17.

U.S. Bureau of the Census, Report WP/98, *World Population Profile: 1998*, U.S. Government Printing Office, Washington, DC, 1999.

Vuuren, Detlef P. and Bouwman, Lex F. Exploring past and future changes in the ecological footprint for world regions. *Ecological Economics,* **52** (2005), pp. 43-62.

World Population Prospects. URL accessed on January 22, 2007. http://esa.un.org/unpp/.

World Population - Wikipedia, the free encyclopedia. URL accessed on January 22, 2007. **http://en.wikipedia.org/wiki/World_population.**