WEB APPLICATION FOR SPATIAL EXPLORATION OF RACIAL DIVERSITY OVER THE ENTIRE UNITED STATES AT 90 M RESOLUTION

*Anna Dmowska, Tomasz F. Stepinski, and Pawel Netzel**

University of Cincinnati, Dept. of Geography, Space Informatics Lab, Cincinnati, OH 45221-0131, USA

1. INTRODUCTION

Geodemography – the study of population types and their distribution in geographical space – provides valuable information for social and health services, economic development and planning. However, geodemographic analysis is rarely conducted on large spatial scale and in high spatial resolution simultaneously as this requires ability to handle "big data" and presents challenges to effective presentation of the results. Here we report on development of SocScape (Social Landscape) – a GeoWeb application designed to explore population density and racial diversity (hereafter referred to simply as "diversity") over the entire conterminous United States at nominal 90 m resolution. SocScape is a computerized map application which works much like the Google Maps and, indeed, includes Google Maps as a spatial reference layer. Demographic layers include maps of population density and diversity calculated for two time points, 1990 and 2000. It enables fast and intuitive exploration of population and diversity patterns starting from the continental scale of the entire U.S. down to the scale of an individual street. Because the demographic layers are given for two different time points separated by a decade SocScape enables precise tracing of spatial change due to population dynamics.

Existing web applications aimed at visualization of spatial distribution of diversity include the Racial Dot Map (http://www.coopercenter.org/demographics) and Mixed Metro (http://mixedmetro.com/). The Racial Dot Map is based on census block data with each person in a block represented as a dot having color corresponding to person's race. It's a neat visualization concept but cannot be utilized for quantitative research. The Mixed Metro shows diversity map similar to that shown in SocScape, but only one state at the time and only at census tract resolution. The Mixed Metro data can be downloaded but not directly from the mapping application. SocScape provides the highest resolution population and diversity map of the U.S.; it provides seamless integration of exploration and data download. The Oak Ridge National Laboratory is developing [1] LandScan USA – a demographic dataset that also covers the entire U.S. at 90 m resolution. However, LandScan USA is not currently available, nor is it expected to be in the public domain once it become available thus limiting its utility to the scientific community.

The data behind SocScape are the gridded demographic datasets covering the entire U.S. and having the cell size of 90 m. The methodology of calculating these grids is described in [2] and briefly outlined in the next section. Computational part of SocScape relies on an open source software GeoServer (http://geoserver.org) that runs on a dedicated server, and its interface is provided via a web browser; most modern web browsers are supported including Chrome, Firefox, Internet Explorer, and Safari. SocScape can be accessed at http://sil.uc.edu/ .

2. DATA AND METHODS

The ultimate source for the gridded demographic datasets used by SocScape is the U.S. Census 1990 and 2000 block-level data. This data is disaggregated into a grid using dasymetric modeling [3] which utilizes 30 m resolution land cover data provided by the National Land Cover Dataset or NLCD (http://www.mrlc.gov/) as an auxiliary variable. There are two caveats that needs elaboration. First, calculating dasymetric model with 90 m resolution from $\sim$8 millions blocks is computationally expensive. Instead, we calculate [2] dasymetric model from already existing $\sim$1 km grid developed by the Socioeconomic Data and Application Center or SEDAC (http://sedac.ciesin.columbia.edu/). SEDAC grids are products of simple areal weighting interpolation from census blocks, with no auxiliary data used. Applying land cover-based dasymetric modeling to SEDAC grids sharpens their spatial resolution and provides a good approximation to a model disaggregated directly from the blocks.

Second, dasymetric models for 1990 and 2000 need to use not only corresponding census data but also concurrent land cover data NLCD 1992 and NLCD 2001, respectively. Unfortunately, NLCDs 1990 and 2001 feature different categories of land cover which would result in incompatible dasymetric models. We addressed this problem by using so-called retrofit [4] land cover datasets – less accurate classifications of land cover categories which were designed especially to provide correspondence between 1992 and 2001 datasets. Thus Soc-

*Also affiliated with Dept. of Climatology and Atmosphere Protection, University of Wroclaw, Poland

Acknowledgments. This work was supported by the University of Cincinnati Space Exploration Institute.

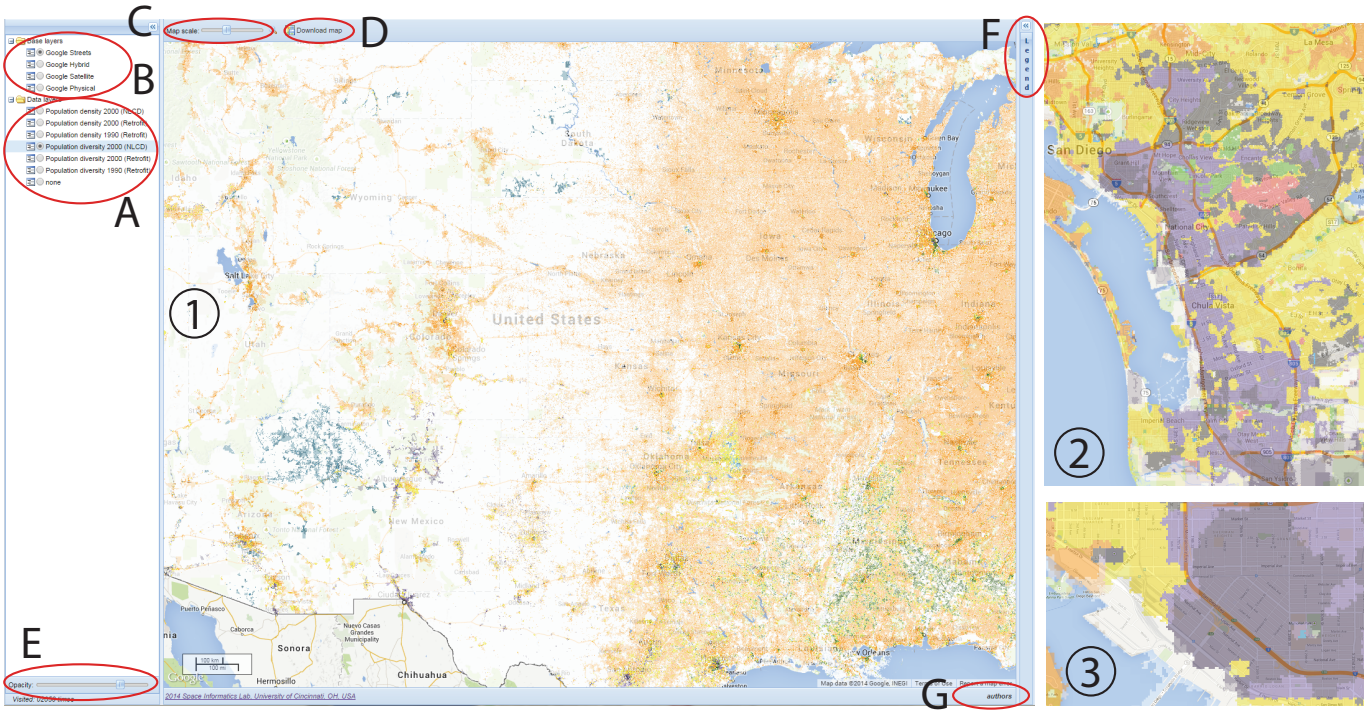

Fig. 1. (1) The screnshot of the SocScape application. Letters A to G lable application features described in the main text. (2) A diversity map overlayed over the Google street map for San Diego area. (3) A diversity map overlayed over the Google street map for the Hilton San Diego Bayfront neighborhood in San Diego.

Scape features six demographic layers, three for population density and three for diversity. Within each category one layer presents the best available dasymetric model based on NLCD 2000, and the remaining two layers present dasymetric models based on retrofit land cover data for 1992 and 2000, respectively. Retrofit-based layer should only be used for assessing 1990-2000 change in demographic patterns. If exploring demographic patterns in 2000 alone the NLCD-based layers should be used.

The demographic layers used in SocScape are stored as integer grids which permits faster response of web application. For population density layers this means that data is classified into 10 bins as shown by the legend. If original, unclassified 90 m resolution population grid is required it can be downloaded from http://sil.uc.edu/ (390 MB). Diversity layers are categorical grids resulting from classification utilizing grids of individual races (white, black, Asian, Hispanic, and others) and the grid of population density. We performed [2] three-dimensional classification based on diversity, race, and population density. With exception of adding population density, our classification follows that of Holloway et al. [5]. There are 33 race-diversity-density categories in the diversity map.

3. USING SOCSCAPE

Panel 1 in Fig. 1 shows a screenshot of SocScape with major features circled in red and labeled by letters A to G. The map layers are listed in the top-left corner and grouped into demographic layers (A) and, for spatial reference, Google Maps layers (B). Upon launching the application only one layer – Google Hybrid – is active. Only one layer from each group can be active at a given time. Demographic layers appear on top of reference layers but their opacity can be adjusted using an opacity slider (E). In order for opacity slider to be activated a demographic layer has to be selected *and* highlighted by clicking the left mouse button on its name.

Map navigation is provided by panning and zooming. In order to pan the map press the left mouse button and drag the map. Use mouse wheel or map scale slider (C) to zoom in and out. Zooming into a selected rectangle is accomplished by pressing Shift button on a keyboard and using mouse to indicate an area into which to zoom in. The georeferenced region of the map that appears in the window can be downloaded (D) in the GeoTIFF format for further analysis providing that the region is smaller than a limit that we have set. If larger region is desired, several smaller regions can be downloaded and put together in using GIS software. Alternatively, the entire (U.S.-wide) maps of population density (390 MB) and diversity (68 MB) can be downloaded at http://sil.uc.edu/downloads.html . The legends to demo-

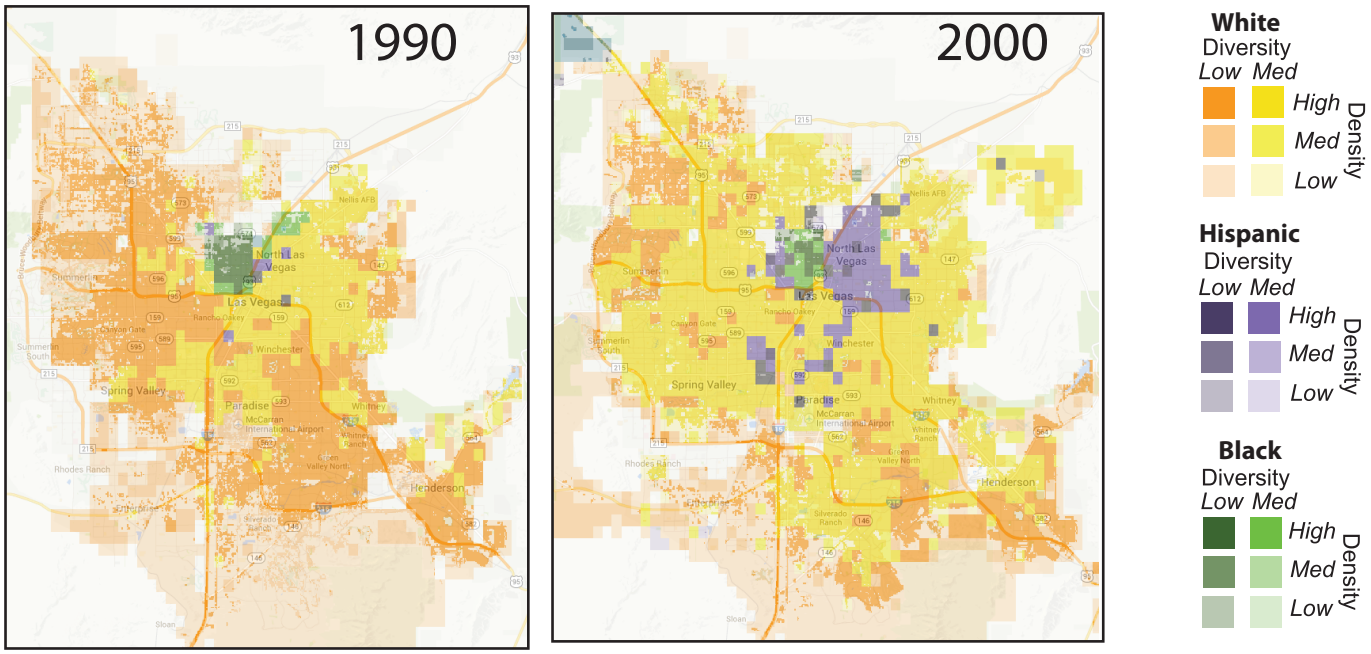

Fig. 2. Using SocScape to explore change in spatial patterns of diversity in las Vegas, Nevada between 1990 (right) and 2000 (left). The legend shows only 18 diversity classes, those most relevant to the Las Vegas region.

graphic maps are hidden but can be revealed by pressing on legend button (F). To make legend to appear permanently the << button located above the legend button needs to be pressed. Only a legend of the selected demographic layer is visible at any given time. As an alternative to using a legend left-clicking the mouse on the map brings up a small window with information about the value of the selected demographic layer at a given point. Clicking on the button labeled "authors" (G) brings up the credits.

Panel 2 in Fig. 1 shows a closeup of the entire map into San Diego, California – the host city of PAA 2015. It shows that, on the basis of 2000 Census, large portions of San Diego area are classified as medium diversity, Hispanic dominated, high population density regions (lighted purple color). Most of the outskirts of San Diego are classified as medium density, white dominated, medium population density regions (yellow color). For meaning of other colors please use SocScape. For definition of the classes see [2]. Panel 3 in Fig. 1 shows a further closeup of the map into an immediate neighborhood of the Hilton San Diego Bayfront where the PAA 2015 will take place. The white color on the map indicates uninhabited areas. It is clear from the panel 3 that diversity at spatial level of individual streets can be explored using the SocScape.

4. DETECTING CHANGE IN DEMOGRAPHIC PATTERNS

SocScape allows analysis of change in spatial pattern of diversity between 1990 and 2000. Fig. 2 illustrates potential for spatio-temporal analysis. It shows diversity map for the city of Las Vegas, Nevada based on 1990 census (left) and 2000 census (right). The diversity pattern has changed significantly. First, the city has expanded. Second, the low diversity, white dominated region (orange) has shrank to only the outer suburbia of the city and has been replaced by the medium diversity, white dominated region (yellow). Third, the explosive growth of Hispanic population resulted in creation of large area classified as medium diversity, Hispanic dominated region (purple). The 1990 and 2000 maps can be downloaded from SocScape and compared quantitatively using GIS software. Because the two grids are co-registered, a pixel-by-pixel change can be analyzed in the same fashion as land cover maps from two different years are compared in the field of remote sensing. Finally, a diversity class transition matrix can be calculated to give the most detailed analysis of spatio-temporal change. Note that such direct comparison of diversity patterns between two years is not possible with diversity maps based on census tracts because boundaries of the tracts change from one census to another.

5. DIVERSITY MAPS: GRIDS VS. TRACTS

High resolution, gridded geodemographic data, like the one used in SocScape, offers number of advantages over traditional data based on census tracts. One advantage has been mentioned in the previous section – gridded data from two different years can be immediately quantitatively compared without any adjustments. Fig. 3 illustrates other advantages of high resolution grids. This figure shows diversity patterns in the region around the city of Fresno, California. The left

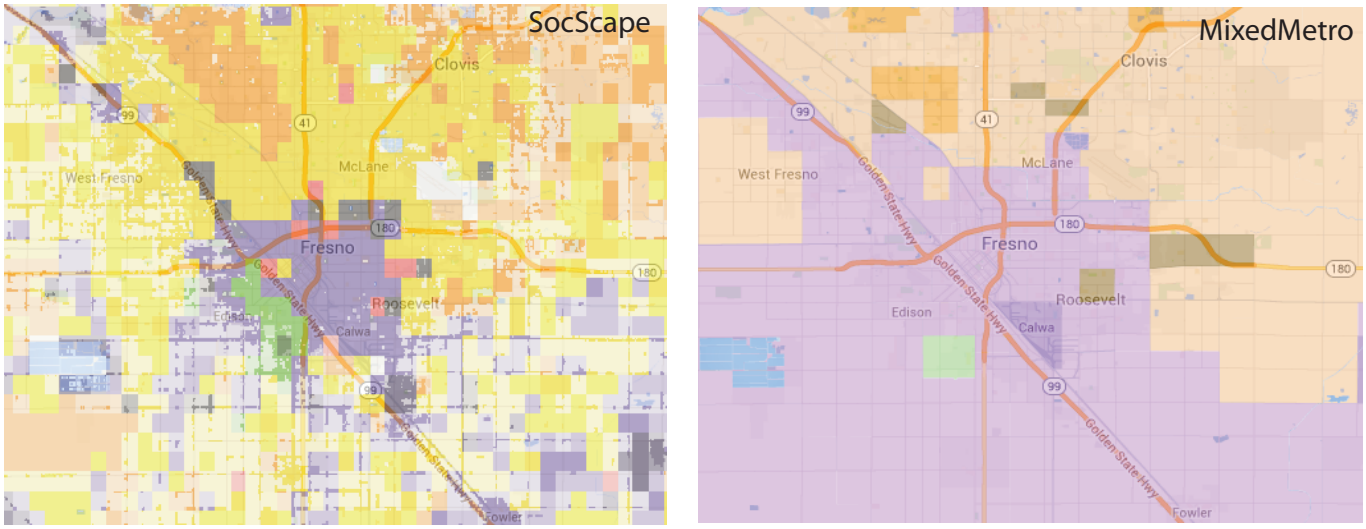

Fig. 3. Comparison of grid-based (right) and tracts-based maps of diversity in the region around the city of Fresno, California. Both maps pertain to 2000 Census.

panel is the SocScape 90m grid-based map and the right panel is the Mixed Metro tracts-based map. The legends to the two maps are not the same but the colors roughly correspond to each other. The tract-based map assigns diversity class (color) to the entire area of the tract even if the significant portion of the tract is uninhabited or very sparsely populated. This is why the Mixed Metro suggests that significant portion of the Fresno region can be classified as medium diversity, Hispanic dominated (light purple) which is incorrect. It also fails to identify concentration of Hispanics near the city center. Grid-based map has also much finer spatial resolution because it uses block data which are further disaggregated into 90m cells; many fine details of diversity pattern seen in SocScape are absent from the Mixed Metro map.

6. CONCLUSIONS

SocScape is a new web-based tool for geodemographic exploration of the U.S. using high resolution gridded data. The gridded data is calculated from the U.S. Census tabulated data using dasymetric modeling. Currently, we have gridded [2] seventeen 1990 and 2000 Census variables pertaining to total population, race and ethnicity, age, and income and use them to construct high resolution maps of population counts and diversity. Future work will concentrate on gridding additional variables, making available additional maps, and on calculating grids for the 2010 Census variables.

7. REFERENCES

[1] B. Bhaduri, E. Bright, P. Coleman, and M. L. Urban, "LandScan USA: a high-resolution geospatial and temporal modeling approach for population distribution and dynamics," *GeoJournal*, vol. 69(1-2), pp. 103–117, 2007.

[2] A. Dmowska and T.F. Stepinski, "High resolution dasymetric model of U.S demographics with application to spatial distribution of racial diversity," *Applied Geography*, vol. 53, pp. 417–426, 2014.

[3] Jeremy Mennis and Torrin Hultgren, "Intelligent Dasymetric Mapping and Its Application to Areal Interpolation," *Cartography and Geographic Information Science*, vol. 33, no. 3, pp. 179–194, Jan. 2006.

[4] J. A. Fry, M. J. Coan, C.G. Homer, D. K. Meyer, and J. D. Wickham, "Completion of the National Land Cover Database (NLCD) 19922001 Land Cover Change Retrofit product u.s. geological survey open-file report 20081379, 18 p.," Tech. Rep., U.S. Geological Survey Open-File Report 20081379, 2009.

[5] S. R. Holloway, R. Wright, and M. Ellis, "The Racially Fragmented City? Neighborhood Racial Segregation and Diversity Jointly Considered," *The Professional Geographer*, vol. 64, pp. 63–82, 2012.