Southwestern Geographer
Vol.4, 2000, pp 16-29

The Relationship between Brightness Values from a Nighttime Satellite Image and Texas County Population

Linda F. Prosperie and J. Ronald Eyton

The Defense Meteorological Satellite Program (DMSP) Operational Linescan System (OLS) acquires nighttime images using a photomultiplier tube sensing system. In addition to recording cloud cover by moonlight and starlight, the sensor images the light emitted from urban and built-up land as well as transients such as oil flares, forest fires, lightning, and the aurora. Attempts to correlate census population counts with pixel brightness are often confounded by an inability to define appropriate urban boundaries. An alternative approach using Texas counties as the observation units produces a strong relationship (r^2=.974) between pixel volume brightness counts and census population counts for an image acquired in 1995. The residuals for the most populated counties in Texas are examined and the reasons for the under-reported and over-reported values are explained in geographical terms. *Key Words: remote sensing, correlation analysis, regression analysis, population.*

Introduction

The Defense Meteorological Satellite Program (DMSP) Operational Linescan System (OLS) produces nighttime images that have transcended beyond the initial meteorological objectives involved with weather forecasting. In addition to recording cloud cover by moonlight and starlight, the photomultiplier tube (PMT) sensor, only used at night, also records the light emitted from urban and built-

Linda Prosperie is a doctoral student in the Department of Geography at Southwest Texas State University, San Marcos, TX 78666. Her email is lp38005@swt.edu. J. Ronald Eyton is Professor in the same department. His email is je09@swt.edu. *The authors' would like to thank the reviewers and the editor for their insightful and helpful comments.*

up land, transients, such as oil flares, forest fires, lightning, and the aurora (Croft 1978, Brandli 1978). The use of imagery for observing city light distributions and for providing the basis for population studies is not new. Tobler's (1969) demonstration that settlement size coefficients could be estimated from satellite photography obtained from the Gemini manned space flight program was the initiating research that involved remote sensing as a means of estimating human population parameters. Ogrosky (1975) found a high degree of association between population and the logarithm of image-areas classified as urban using aerial high-altitude color-infrared transparencies. Lo and Welch (1977) demonstrated how Chinese city population counts could be estimated from the measurement of built-up urban areas on the LANDSAT satellite images. Forster (1983) examined some of the relationships between socioeconomic variables and satellite spectral signatures of an urban area using LANDSAT data acquired over the Sydney, Australia metropolitan area. Multiple regression analysis was used to develop predictive equations for surface-cover proportions, housing density, relative average housing values, and a residential quality index.

In 1980, two studies modeled urban population trends and urban energy utilization patterns using census data, urban energy consumption statistics, and satellite imagery. Welch 1980) found striking correlations between census data populations for cities throughout China and the area of these cities delineated on LANDSAT images. In addition, the author demonstrated high correlations between urban energy consumption and the volume of "illuminated urban areas" (IUAs) recorded from DMSP/OLS images. Welch and Zupko (1980) investigated urbanized area energy utilization patterns for thirty-five U.S. cities and also found high correlations between population, urban energy utilization, and the brightness distribution of IUAs. Imhoff et al. (1997) explored the use of nighttime DMSP/OLS images of city lights to estimate the impact of urban land use on soil resources in the U.S. They found that nighttime images of the Earth can be transformed, with appropriate temporal thresholds for reducing noise and "blooming" (or excursion of light into water bodies) of the city lights around coastlines, to produce populated land cover classes. The authors suggested that when calibrated regionally to account for housing and population characteristics, the DMSP/OLS data set may even be a substitute for census data especially in lesser developed countries where census data are scarce, often in error, and infrequently collected.

Elvidge et al. (1997b) produced a global inventory of human settlements using DMSP/OLS images of Europe and South America. Examination of the relationship between area lit, population, and energy-related carbon emissions for fifty-two countries revealed substantial variations between individual countries possibly due to cultural preferences and differences in the sources of electric power. Sutton (1997) presented a means of modeling human population density of urban clusters identified within the continental U.S. using DMSP/OLS nighttime satellite imagery and a Geographic Information System (GIS). The author's method defined the edges of the urban areas rather than attempting identification of the center of an urban cluster. The results demonstrated the use of nighttime satellite imagery as a proxy measurement of urban extent. The DMSP/OLS was also used by Elvidge et al. (1997a) for mapping populated areas using 236 images acquired between October 1994 and March 1995 for the U.S. The authors described the unique characteristics of the OLS for observing faint sources of visible and near-infrared (VNIR) emissions and indicated how time-series analysis can be used to discriminate between stable VNIR emission sources such as cities, towns and industrial sites and ephemeral VNIR emissions such as fires and lightning. The authors suggested that a global map of city lights available to the scientific community would aid in the analysis of social, environmental, and energy issues such as the detection of blackouts and brownouts in the electrical grid or the expansion of urban areas.

Sutton et al. (1997) compared nighttime satellite imagery and population density for the continental U.S. and concluded that DMSP/OLS imagery "does not show a strong simple quantitative correlation with human population density." The authors investigated the correlation between population density and DMSP/OLS brightness values by aggregating DMSP/OLS light values to state and county levels. The sum of each county's DMSP/OLS pixel values was used as the independent variable and county population counts as the dependent variable in a linear regression model that produced an r^2 of .50. This research, however, shows that, for at least the state of Texas, a strong curvilinear relationship does exist between the image brightness values and population.

Three research objectives were carried out for this study. First we determined the relationships between Texas county population and DMSP/OLS pixel area brightness counts and pixel volume brightness counts. Second we explicated the relationships found in the context of the methodology. And third, we examined the residuals of our analysis for the most populated counties in Texas and explain in geographical terms the reasons for under-reported or over-reported values.

Data

DMSP/OLS imagery has been acquired from 1972 to the present through the operation of satellite sensors capable of detecting the reflections and emissions from cities and towns and the global distribution of clouds and cloud-top temperatures. Images are obtained for a 3000-kilometer-wide scan twice each twenty-four hour period, one scan during daylight and a second scan at night. The OLS is an oscillating scan radiometer with two spectral bands, visible (VIS) and thermal infrared (TIR). The OLS VIS band signal is intensified at night using a photo multiplier tube (PMT) to acquire the nighttime images at a nadir pixel resolution of 2.7 kilometers (km). This sensor is capable of measuring emitted radiation between 0.47 – 0.95 micrometers (mm) and can detect city lights, fires (including gas flares), lightning, and the aurora. Croft (1978), Welch and Zupko (1980), and Elvidge et al. (1997b) provide detailed explanations of the characteristics of the DMSP/OLS.

The nighttime image covering the east and central portions of the U.S. was acquired February 1, 1995 (Figure 1). Texas appears in the lower left and the New York megalopolis region is clearly visible in the upper right. Towns with populations as low as 1477, such as Comfort, Texas, are visible in the scene. Other towns with similar populations may not appear depending on whether the light distribution is concentrated (as in a business district) or spread out along a major route that runs through town. Commercial operations, such as storage facilities, cement plants, and factory outlet shopping malls, may be the principal (and sometimes only) light source. The 1257 scan lines by 1404 pixels per scan line nighttime image was downloaded from the DMSP/OLS archive (United States Department of Commerce 1998) in Graphic Interchange Format (GIF) at 8 bits per pixel. In this format, each pixel value ranges from 0 to 255 rather than 0 to 63 as in the original DMSP/OLS imagery (Elvidge et al. 1997b, Sutton et al. 1997).

A subset of the U.S. DMSP/OLS image containing Texas, not registered or geometrically corrected, was extracted and rotated to produce a north-oriented central longitude. The DMSP/OLS subset image was then overlaid with 254 Texas county borders (Figure 2). County outlines were created from latitude and longitude coordinates extracted from the Census Bureau's Topologically Integrated Geographic Encoding and Referencing (TIGER) database of geographic information (U.S. Bureau of the Census 1998). The latitude-and-longitude-delineated county polygons were converted from vector to raster format then transformed to raster rows and column coordinates in register with the Texas DMSP/OLS image. Two dozen points, distributed as evenly as possibly across the image, were used to register the Texas DMSP/OLS image to

Figure 1. U.S.A. DMSP/OLS image acquired February 1, 1995.

the county outlines. Where no lights existed in the corners of the state, as in the cases of Big Bend National Park in the southwest and the rural portions of the northwest panhandle, registration points outside of Texas were used. The registration error was reduced to near the resolution of the image, 2.7 km, by using smaller cities in the registration process.

The "blooming" effect, encountered by Imhoff et al. (1997) using a data set of the entire U.S., was not an issue in this study. The most brightly illuminated areas along the Texas coastline do not demonstrate a problem with blooming. Freeport and Corpus Christi are not extraordinary light sources and Houston, bounded by Galveston Bay, retains a coherently lit polygon for Harris County. These areas do not "bloom" as do large cities (e.g. Chicago) and megalopolitan cities (such as those

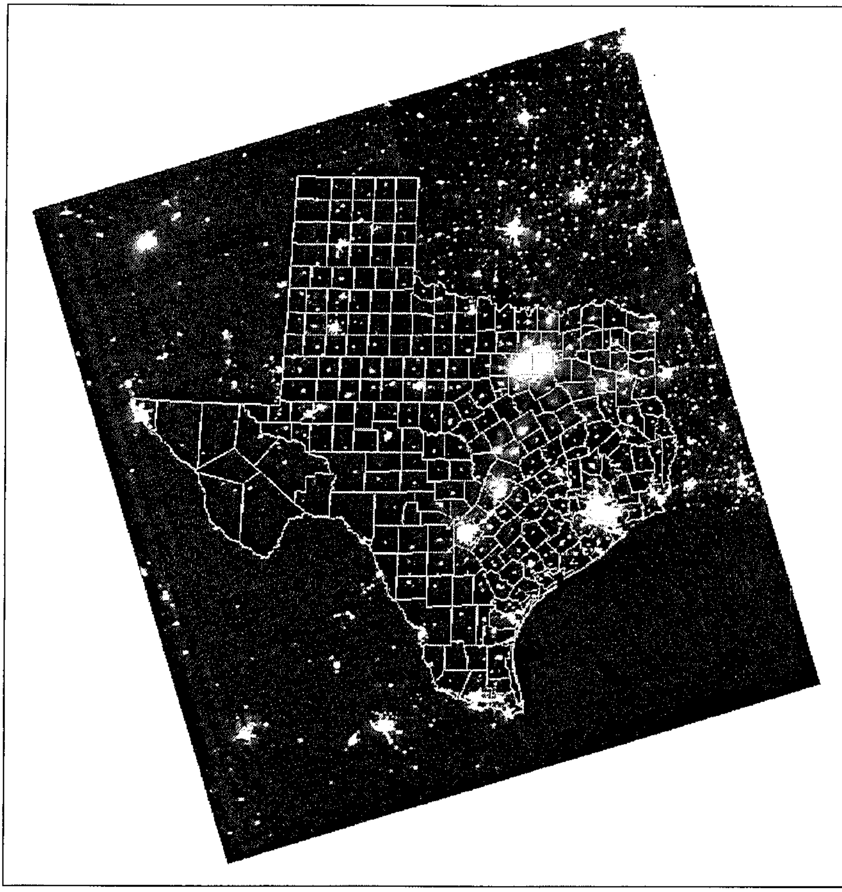

Figure 2. DMSP/OLS subset image (February 1, 1995).

found along the Florida coast and the eastern seaboard).

Population data used in this study were obtained from the Texas State Data Center at Texas A&M University (1998). Population estimates from 1995 were used to match the image acquisition date. County-level data were chosen due to the problems associated with defining meaningful urban boundaries and obtaining the matching population counts. Counties as observation units provide an effective and accurate alternative because bright, populated pixels are easy to count within the county border and county populations are reported by the census. Terra Firma (Eyton 1998) software was used throughout this study.

Extracting Brightness Measures

Two brightness pixel counts were obtained systematically from the DMSP/OLS image subset of Texas using a mask comprising the 254 counties (Figure 3). The county mask was created using the county-outline data set in which all of the pixels for county #1 (Anderson) contained a value of 1, all of the pixels for county #2 (Andrews) contained a value of 2, and so on. Because Texas has 254 counties, the mask data set consists of pixel values 1 – 254 for the counties, pixel values of 0 for the areas outside of Texas, and pixel values of 255 for the county boundaries. The mask was used to obtain a count of the bright pixels in each county. A threshold gray-level value (GLV) determined the separation of bright pixels from background or dark pixels. Twelve samples of small towns with minimal brightness were selected and the surrounding pixel values examined, resulting in selection of a GLV of 23 as the break between background pixels and lit pixels. The frequency of pixels equal to or greater than the threshold value of 23 were counted for each area (area count) and provided a measure of the county area that is considered "bright." Another count (volume count) summed the GLVs of the pixels that are equal to or greater than the threshold value of 23 to provide an indication not only of area but an integrated county "brightness."

Correlation, Regression, and Residual Analysis

The results of regressions performed (Table 1) are compared to the results of other authors (Table 2). The best model using DMSP/OLS imagery shows a strong curvilinear relationship (r^2 = 0.974) between

1. Create the county mask from the DMSP/OLS subset image:
   - For all pixels in the subset image:

     If the subset image pixel is outside Texas, mask pixel value = 0
     If the subset image pixel is within county #1, mask pixel value =1
     If the subset image pixel is within county #2, mask pixel value = 2

     If the subset image pixel is within county #254, pask pixel value = 254.

     Note: Boundary pixels were only used in a mask to make the images, but were not used in the brightness counts. For example, county 253 pixels abutted county 254 pixels; no pixels were used to define the boundary between them.

2. Determine the minimum brightness threshold value:
   - Visually select twelve areas of minimum brightness from the subset image
   - Inspect the GLVs of the selected areas
   - Choose a common break from the twelve areas between background pixels and lit pixels

3. Compute the area and volume brightness measures for each county.
   - For all pixels in the subset image:

     If the subset image pixel value is greater than or equal to the minimum brightness threshold of 23:
     area brightness count, = area brightness count, +1
     volume brightness county, = volume brightness count, + subset image pixel value

     where i = the mask pixel value or county index.

Figure 3. Steps for extracting brightness measures.

population and volume count. However, the relationship between Texas county population and DMSP/OLS brightness counts demonstrates an increasing rate of change as described by the curvilinear model. As a city grows in size, brightness also grows at an increasing rate. Cities that are small tend to have infrastructures that only serve local needs. Larger cities may have infrastructures such as commercial and industrial complexes that serve regional, national, and international markets. This change in the infrastructural focus for larger cities may, in part, account for the disproportionate increase in brightness. Other considerations such as the "stacking" of population in high-rise structures in larger populated urban

Table 1. Regression analysis r^2 values for Texas.

Population vs. Area Count (linear)	0.866
Population vs. Volume Count (linear)	0.884
Population vs. Area Count (quadratic)	0.876
Population vs. Volume Count (quadratic)	0.974
Population (log) vs. Area Count (log)	0.773

Table 2. Regression analysis r^2 values from other studies.

Author	Observations	Model	r^2	
			East	West
Welch and Zupko (1980)	35 samples across the U.S.	Population vs. Energy Use	0.77	0.92
		Population vs. IUA Volume	0.92	0.90
		Energy utilization vs. IUA Volume	0.79	0.86
Ogrosky (1975)	18 samples across the Puget Sound	Population vs. Area (log)	0.964	
Sutton et al. (1997)	5000 samples across the U.S.	Population vs. Area		
		State level	0.69	
		County level	0.50	
Welch (1980)	124 cities in China (LANDSAT)	Population vs. Area	0.75	
	13 large cities in China (LANDSAT)	Population vs. Area	0.82	
	18 IUAs in China (DMSP/OLS)	Energy Consumption vs. Volume	0.89	
Sutton (1997)	Continental coverage of the U.S. at a resolution of one square kilometer	Cluster Population vs. Cluster Area (ln)	0.975	
Lo and Welch (1977)	13 mainland cities in China	Population vs. Area (log)	0.56	

areas may also contribute to the curvilinear form of the relationships.

The residuals and the quadratic regression line for population versus brightness volume count ($r^2 = 0.974$) exhibit a systematic pattern (Figure 4). All counties with a positive residual are under-reported by the regression model and appear above the regression line. The cities (Dallas, San Antonio, El Paso, and Plano) found within each county appear not to have enough brightness for its population and can be described as "compact" urban and built-up lands. Those counties with negative residuals, below the line, show the opposite relationship. The cities (Houston, Fort Worth, McAllen, Edinburg, Denton, Freeport, and Lake Jackson) within these counties are too bright for the population and hence are over-reported by the regression model. These cities can be described as "spread out" and have more per capita street lighting associated with the diffused population. Some counties, such as Brazoria, contain well-lit large industrial complexes (oil refineries) and cause over-reporting by the regression model. Other counties like Hidalgo contain a number of smaller cities (McAllen and Edinburg) that create a more "spread out" infrastructure, particularly along the highway connecting them. This observed systematic pattern in the residuals meets our expectation determined from the characteristics of cities within the counties. In this respect we agree with the suggestion that "DMSP imagery underestimates the population density of urban (compact) centers and overestimates the population density of suburban (spread out) areas" (Sutton's et al. 1997).

Conclusions

The relationship between Texas county population and DMSP/OLS brightness volume count is much stronger than suggested by Sutton et al. (1997). Larger cities produce greater errors in the regression analysis and in the case of Texas, the counties containing the largest cities (Hous-

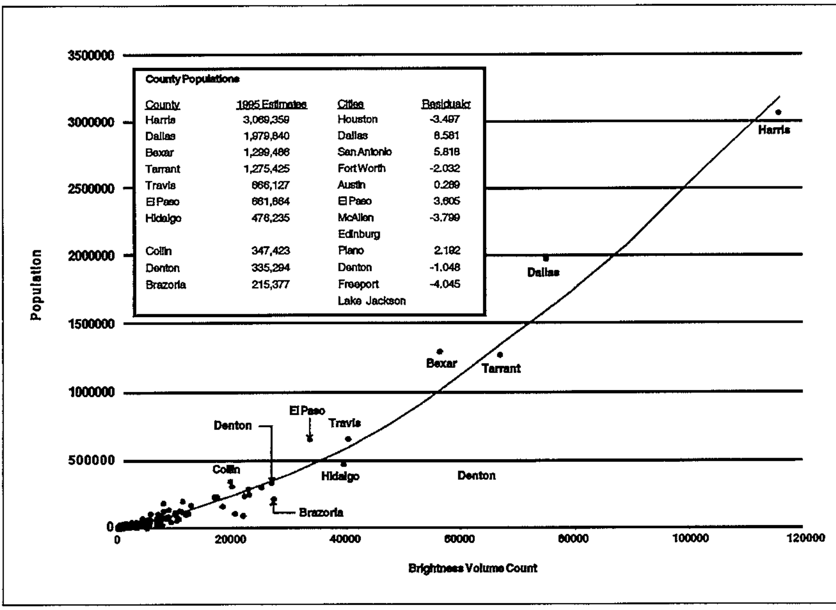

Figure 4. Scatter diagram with the quadratic regression line for population versus brightness volume count.

ton, Dallas, Fort Worth, San Antonio, Austin, El Paso) do not confound the relationship as readily as New York, Los Angeles, and Chicago do at the national scale. The use of counties to aggregate both the population and the volume brightness count appears to involve the appropriate observation unit and eliminates the problem of identifying the boundaries of urban places. The relationship, at least at the state level, appears to be curvilinear rather than logarithmic. Excluding the background pixels not associated with populated areas strengthens the relationship. This exclusion of essentially unpopulated areas at a county level was not done in other studies and may be the principal reason for the high r^2 value obtained for the 254 Texas counties in this study. Analysis of residuals shows that for the counties with large populations, the distribution of people and associated lighting depends on the character of the urban and built-up land in terms of how the city has grown.

References

Brandli, H. W. 1978. The Night Eye in the Sky. *Photogrammetric Engineering and Remote Sensing.* 44(4):503–505.

Croft, Thomas A. 1978. Nighttime Images of the Earth from Space. *Scientific American.* 239:86-98.

Elvidge, C.D., K.E. Baugh, V.R. Hobson, E.A. Kihn, H.W. Kroehl, E.R. Davis, and D. Cocero. 1997a. Satellite Inventory of Human Settlements Using Nocturnal Radiation Emissions: A Contribution for the Global Toolchest. *Global-Change-Biology.* 3(5):387-395.

Elvidge, C. D., K. E. Baugh, E. A. Kihn, H. W. Kroehl, and E. R. Davis. 1997b. Mapping City Lights With Nighttime Data from the DMSP Operational Linescan System. *Photogrammetric Engineering and Remote Sensing.* 63(6):727–734.

Eyton, J. R. 1998. *Terra Firma,* unpublished software package for digital image processing and digital terrain modeling, Southwest Texas State University.

Forster, B. C. 1983. Some Urban Measurements from Landsat Data. *Photogrammetric Engineering and Remote Sensing.* 49(12):1693-1707.

Imhoff, M. L., W. T. Lawrence, C. D. Elvidge, T. Paul, E. Levine, M.V. Privalsky, and V. Brown. 1997. Using Nighttime DMSP/OLS Images of City Lights to Estimate the Impact of Urban Land Use on Soil Resources in the United States. *Remote Sensing of Environment.* 59:105–117.

Lo, C.P. and R. Welch. 1977. Chinese Urban Population Estimates. *Annals of the Association of American Geographers.* 67(2):246-253.

Ogrosky, Charles E. 1975. Population Estimates from Satellite Imagery. *Photogrammetric Engineering and Remote Sensing.* 41:707-712.

Sutton, P. 1997. Modeling Population Density with Night-time Satellite Imagery and GIS. *Computers, Environment and Urban Systems.* 21(3/4):227–244.

Sutton, P., D. Roberts, C. Elvidge, and H. Meij. 1997. A Comparison of

Nighttime Satellite Imagery and Population Density for the Continental United States. *Photogrammetric Engineering and Remote Sensing.* 63(11):1303–1313.

Texas State Data Center at Texas A&M University. Total Population Estimates for Texas Counties and Places. [Online] Available http://txsdc.tamu.edu/txpop95.html, July 7, 1998.

Tobler, W.R. 1969. Satellite Confirmation of Settlement Size Coefficients. *Area.* 1:30-34.

United States Bureau of the Census. 1994. *County and City Data Book.* Washington, D.C.: U.S.A. Government Printing Office.

United States Bureau of the Census. County Boundary Files. [Online] Available http://www.census.gov/tiger/boundary/county/, August, 1998.

United States Department of Commerce, National Oceanic & Atmospheric Administration, National Environmental Satellite, Data and Information Service, National Geophysical Data Center. Nighttime Lights of the Continental United States. [Online] Available http://spidr.ngdc.noaa.gov:8080/production/image/BIOMASS/usa_lights.gif, July, 1998.

Welch, R. 1980. Monitoring Urban Population and Energy Utilization Patterns from Satellite Data. *Remote Sensing of Environment.* 9:1-9.

Welch, R. and S. Zupko. 1980. Urbanized Area Energy Patterns from DMSP Data. *Photogrammetric Engineering and Remote Sensing.* 46(2):201-207.