Building and Evaluating Models to Estimate Ambient Population Density

Paul C. Sutton, Chris Elvidge, and Tom Obremski

Abstract
The concept of human population density is quite simple: the number of persons occupying a given area. Nonetheless, practical representations of population density must use appropriate spatial and temporal scales of measurement to be useful. The 11 September 2001 attack on the World Trade Center in New York is a poignant example: "How many people were in the two World Trade Center buildings at 0830 local time?" Population density data derived from most-national censuses is a residential measure of population density and consequently does not capture non-residential population density. Human mobility suggests that a non-residential or ambient measure of population density may be a more useful representation for some applications. Ambient population density in this sense is a temporally averaged measure of population density that takes into account where people work, sleep, eat, drive, shop, etc. Short of implanting a GPS reciever into everyone's skull and tracking their spatio-temporal behavior, it is extremely difficult to make direct measurements of ambient population density. This paper explores some theoretical and empirical efforts at estimating ambient population density and proposes a quantitative means for evaluating their validity. The three models of population density examined are LandScan, Gridded Population of the World (GPW), and a simple empirical model derived from nighttime satellite imagery provided by the Defense Meteorological Satellite Program's Operational Linescan System (DMSP OLS). These measures are compared to both residential and employment-based measures of population density in the Los Angeles metropolitan area. The GPW, LandScan, and DMSP OLS models of ambient population density described here all make foundational contributions to future efforts at filling the gap in social, economic, and demographic information for parts of the world where such data are unavailable. The proxy measures of population density described here show promise for many applications, including improved mapping of population distribution and as a supplement to census enumerations in many parts of the world.

Introduction
Population density is a commonly used characterization of geographic space at a wide range of scales. Measures of population density are useful in hazard planning and response, environmental impact assessment, transportation planning, economic decision-making, and numerous other applications. Useful measures of population density must be made at appropriate, application specific, spatial and temporal scales. Nonetheless, population data are not always available at the spatial and temporal resolutions required. This paper explores some empirical and theoretical approaches to estimating ambient population density. Ambient population density is a temporally averaged measure of population density that incorporates human mobility. The concept of "ambient population density" was suggested by Dobson *et al.* (1999) in a paper describing the production of a global population dataset called LandScan. Representations of ambient population density may be more useful for certain applications than are existing traditional measures of population density.

The concept of population density can be elusive when one attempts to characterize it across spatial and temporal scales. One can concievably conduct the following thought experiment to gain a sense of the complexities associated with representing population density at varying spatio-temporal scales: At some fine spatial scale, population density becomes binary, e.g., one person per unit area (with perhaps the exception of crowded multi-story buildings). A binary map of population density at such a scale can be imagined with dots representing individuals.

A map of this nature is a snapshot in time. Yet people move. To represent this motion, an animated map of moving dots could portray a binary scale population density map through time. Imagine a pixel or a small fixed square location in this animated map. Sometimes it would contain a person and have a value of one and other times it would be empty with value zero. An average value somewhere between one and zero could be calculated for this pixel or rectangular space as time progressed through seconds, minutes, hours, days, weeks, months, or years. It is quite likely that this average population density of this cell would be quite variable over the course of a day or perhaps even a week, yet at some point in time the average would settle down and not vary significantly. The time at which the average population density of this cell or pixel stopped varying might be an appropriate temporal scale for measuring the ambient population density of this pixel. At this point it seems appropriate to shift from a pointillist perspective to pixellated perspective.

What is the appropriate method to characterize the population density of a pixel in time and space? Would all pixels at this spatial scale stabilize in population density over the same period of time? A pixel in a crosswalk at a busy intersection in a city might reach a stable average population density in the course of a day, yet it might take weeks or months for a pixel in a conference center to stabilize due to more extreme fluctuations in the flow of people

Paul C. Sutton is with the Department of Geography, University of Denver, Denver, CO 80208 (psutton@du.edu).

Chris Elvidge is with the NOAA National Geophysical Data Center, 325 Broadway, Boulder, CO 80303 (Chris.Elvidge@noaa.gov).

Tom Obremski is with Statistics and Operations Technology, Daniels College of Business, University of Denver, Denver, CO 80208 (tobremsk@du.edu).

Photogrammetric Engineering & Remote Sensing
Vol. 69, No. 5, May 2003, pp. 545–553.

0099-1112/02/6905–545$3.00/0

through it. Pixels in a national park might not stabilize until years have passed due to seasonal variation in the flow of people through them. In any case, temporal variability is an issue to account for in conceptions of population density.

The spatial variability of population density also raises interesting questions. Just as we smoothed the map through temporal averaging, this map can be smoothed with spatial averaging. Spatial averaging could be accomplished by increasing the size of the pixels to larger and larger sizes. As this is done pixels will take on a greater range of values between zero and some number that is a function of the size of the pixel. As this smoothing increased and successive representations of it were presented to an observer, one would begin to see both increasing variation in levels of population density and increasing smoothness of population density. However, as the scale became increasingly coarse, this smoothness would begin to disappear and the variability of the population density would actually decrease. If this exercise were performed on a dataset of the continental United States and the pixels were as large as typical urban centers, the map would become almost binary again. The pixels over urban centers would have large values which varied according to the average population density of the urban centers they represented, and the rest would have slightly varying low values. In any case, the spatial and temporal variability of population density presents interesting problems of definition and measurement. The complexities associated with human spatio-temporal behavior are beyond the scope of this paper as are the intractable problems associated with scale and aggregation. However, the simple empirical investigations described here may shed light on theoritical, empirical, and practical attempts at measuring and representing population density using remotely sensed imagery and geographic information systems.

Data and Methods

The three models of population density used in this study are the Gridded Population of the World (GPW) (Tobler *et al.*, 1997), LandScan (Dobson *et al.*, 2000), and a model derived from nighttime satellite imagery provided by the Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS) (Sutton 1997; Elvidge *et al.*, 1998). The datasets used to evaluate these estimates of ambient population density were a 1-km^2 grid of population density derived from 1990 Census block groups of the conterminous United States and a Public Use Micro Station (PUMS) dataset of employment in the Los Angeles area. A brief description of these datasets follows.

Gridded Population of the World (GPW): A Single Variable Model

The Gridded Population of the World (GPW) started as The Global Demography Project, a joint effort of the National Center for Geographic Information and Analysis (NCGIA), the Center for International Earth Science Information Network (CIESIN), and Environmental Systems Research Institute (ESRI)(Tobler, 1995). The Global Demography Project produced a gridded global representation of the world's population circa 1994. Population counts were obtained from best available first- or second-level sub-national population counts (about 16,000 administrative units), extrapolated to 1994 populations, and interpolated to a 5- by 5-minute grid using a smoothing algorithm developed by Tobler (1979). This smoothing eliminates arbitrary "steps" in population density that occur at administrative boundaries.It is the only manipulation of the data that has the potential to capture ambient population density. GPW is essentially a measure of residential rather than ambient population density. GPW is considered a single variable model because the only information used to derive the values is population counts from vector maps derived from official census figures. In some cases the population count of the polygons is extrapolated to 1990 or 1995 from the most recent available census; however, the population counts are the only information used to produce gridded population estimates. LandScan, on the other hand, uses a diverse assortment of spatial datasets in addition to census data to produce an estimated global grid of population.

The LandScan Global Population Project: A Multiple Variable Model

The LandScan gridded global population dataset was developed at Oak Ridge National Laboratories to provide population data of fine enough spatial resolution to be useful in preparing for and responding to natural and manmade disasters (Dobson *et al.*, 2000). The cell resolution of LandScan is 1 km^2; this is 100 times finer a spatial resolution than GPW. The creators of LandScan coined the phrase "ambient population density" and explicitly built their database to reflect ambient as opposed to residential population density. LandScan is a multiple variable model in that it uses many input variables such as roads, topographic slope, land cover, populated places, nighttime lights, exclusion areas, urban density factors, and coastlines to estimate ambient population density. LandScan is spatially interpolated to a finer resolution than the actual census values by applying many theories about the concentration of people near roads, coastlines, and urban areas. This borrows from theory about improved population allocation modeling by utilizing appropriate ancillary information and the capabilities of geographic information systems (Landford and Unwin, 1994; Deichmann, 1996). The parameters that control the probability models for allocating population in the LandScan model are spatially variable (i.e., they can vary from province to province) and empirically derived. Landscan is similar to GPW in that it uses best available population data from polygonal maps that varied from coarse resolution (state or province level) in some countries to fine resolution (tract level) in other countries. LandScan's cell values aggregate to these polygon values as do the GPW products; however, the spatial variability of population density within these polygons can be higher in LandScan because of the use of ancillary data.

Nighttime Satellite Imagery as a Proxy Measure of Ambient Population Density

Another model estimating ambient population density was derived from a stable nighttime lights dataset developed by Elvidge *et al.* (1998). This image is a composite of hundreds of orbits of the DMSP OLS in which the images were screened for clouds and ephemeral lights such as fires and gas flares. Another dataset derived from the DMSP OLS data was a binary image of the urban clusters. This image was used as a mask for separating urban and rural populations and as a means of estimating the population of urban clusters (see Sutton *et al.* (2001) for a description of how the imagery was used to estimate the population of all the urban clusters of the world). The estimate of urban population density at each non-zero pixel of the nighttime satellite image is based on the following model:

$$\begin{matrix}\text{Estimated}\\ \text{Population}\\ \text{Density}\end{matrix} = \begin{pmatrix}\text{Estimated}\\ \text{Cluster}\\ \text{Population}\end{pmatrix} \times \frac{\text{DMSP Low-Gain Pixel Value}}{\text{Sum of all DMSP Low-Gain PixelValues in Cluster}}$$

or

$$E_j = E_k{}^*(lgv_i/\Sigma(lgv_i)) \quad i \in M_k \tag{1}$$

where lgv_i is the low-gain value of pixel i, k is the index of cluster, j and i are the indices of pixels, E_k is the estimated total population of cluster k, E_j is the estimated population density of pixel i, and M_k is the pixel membership of cluster k. A graphical depiction of the above model is provided in Figure 1. The total cluster population used in the model can be actual values or estimates derived from linear regressions described in Sutton *et al.* (2001). This model only estimates ambient population density in urban areas, whereas LandScan estimates ambient population density on all land surfaces.

Data Used for Model Evaluation

The population density standards used for validation of these models were derived from a grid of population density created from the block-group administrative boundaries of the 1990 U.S. census. The square cells in this grid were allocated population density values on a proportional-by-area basis from the block-group polygons of the TIGER files. These reference data were manipulated in three ways to capture human mobility and approximate ambient population density: (1) spatial aggregation to larger pixels, (2) the use of mean filters, and (3) the use of both employment and residence-based population density datasets. The employment-based population density dataset was derived from PUMS data for the Los Angeles area.

Manipulations one and two are isotropic, whereas in reality there will be regionally varying anisotropic mobility patterns that will influence any accurate representation of ambient as opposed to residence-based population density (e.g., the urban periphery moving into downtown cores during the daytime). The third manipulation simply took the average of employment- and residence-based measures of population density. This manipulation is empirical and captures some of the anisotropic nature of the difference between residence-based and ambient population density.

Spatial autocorrelation is one attribute of the population density dataset that presents problems for developing accurate models to estimate its high degree of spatial variability, as shown in a correlogram of the population density dataset (Figure 2). The figure shows how a one-pixel mis-registration of a perfect model to the actual dataset will only result in an R^2 of 0.74. This problem is mitigated using spatial aggregation or smoothing with a mean filter.

Spatial aggregation reduces the variability of pixel values towards their mean. Typically, the pixel values have high variability at fine resolutions and lower variability at coarser resolutions; continued aggregation ultimately results in one average population density for the whole planet. Smoothing retains the spatial resolution of the dataset but reduces the variability of the pixels locally (Holloway,

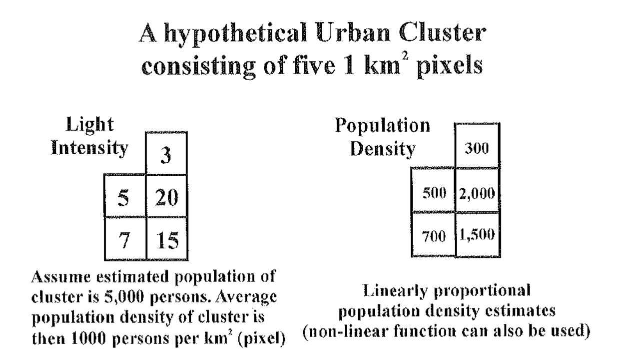

Figure 1. Graphical representation of population density model derived from DMSP OLS imagery.

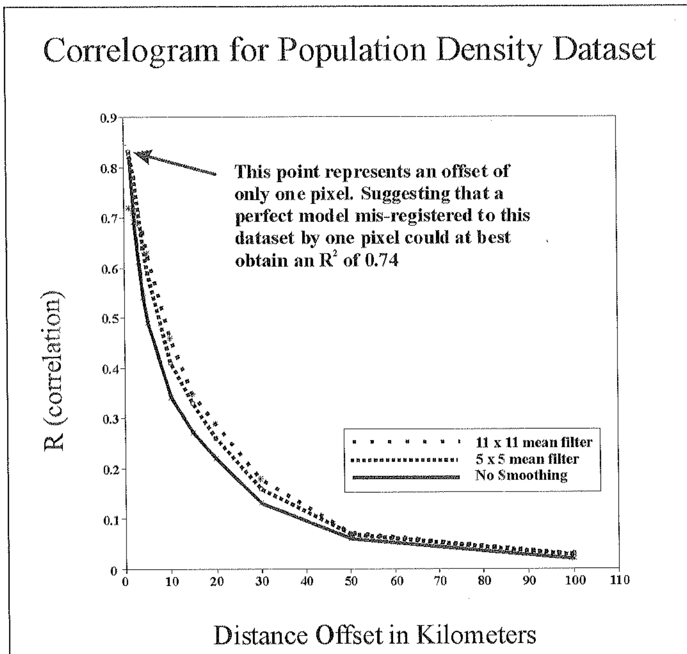

Figure 2. Spatial autocorrelation of population density image.

1958). The correlogram for a smoothed population density dataset does not drop off as sharply as shown in Figure 2. In this study the ground truth dataset was smoothed using a 5- by 5-pixel mean filter and an 11- by 11-pixel mean filter. The justification for mean smoothing is that the census data records where people are at night. A mean spatial filter is a method of performing a smoothing that attempts to capture the reality of people moving from home to office to store to school. In the United States in 1975 the average distance commuted from home to work was approximately 9 miles (Long and Boertlein, 1976). Mean filtering is one means of approximating ambient population density that results from human mobility.

The use of a smoothing filter serves two purposes. First, smoothing may produce an image that is more representative of ambient population density through time. Second, smoothing increases the spatial autocorrelation of the data, in essence, reducing the steepness of the decays shown in Figure 2. An area that exemplifies the problem of high spatial variability of residence-based measures of population density is the University of California at Santa Barbara (UCSB) campus and the nearby student community of Isla Vista. The census records show UCSB as having a very low population density, whereas Isla Vista has some of the most densely populated census tracts west of the Mississippi. A mean smoothing of the population density image increases the population density of areas like the UCSB campus, particularly where they border Isla Vista. In addition, the Isla Vista community has low population density commercial districts next to high density residential districts that are resolved by the census block polygons and the resulting 1-km² grid derived from them. This is a good example of high spatial variability at the fine scale. The smoothing filter produces a representation of population density in Isla Vista that spreads the population more evenly between the commercial and residential areas in the community.

The model of population density derived from nighttime satellite imagery was evaluated on the 1-km² population density dataset, and on aggregations of the dataset to pixels with 5- and 10-km^2 sides. The larger pixel images were simply mean aggregations of the finer resolution image. The resulting "predictions" of population density were then compared to both the smoothed and un-smoothed "reference database" population density.

The third method for evaluating the GPW-, LandScan-, and Nighttime-Image-based estimates of population density focused on the Los Angeles metropolitan area. The "reference database" of ambient population density used for the Los Angeles metro area was a simple arithmetic average of the residence-based measure of population density derived from 1990 census data and an employment-based measure of population density derived from PUMS data.

Results and Analysis

The urban cluster containing Minneapolis and St. Paul provides a good visualization of the appearance of these models and several of the standards to which they were compared (Figure 3). The images in Figure 3 include the unsmoothed population density derived from the block group polygons of the 1990 U.S. census, the 5- by 5-pixel and 11-by 11-pixel mean filtered versions of those data, the model derived from the DMSP OLS imagery, and the 5- by 5-minute global demography project. Because the models derived from the DMSP OLS nighttime imagery are only produced over urban areas, the results reported in Table 1 are limited to lit areas in the DMSP OLS imagery. By focusing on these pixels, we analyze the 10 percent of the land which contains almost 80 percent of the people. The continental United States had 5,881 urban clusters with an average population of 32,952. This represents 193.8 million of 239.4 million in the dataset for 1990, equal to 79 percent of the total population. As an estimate of urban population, it is 4 percent off the Population Reference Bureau's value of 75 percent (Haub and Cornelius, 1998). The models were compared at 1-, 5-, and 11-km^2 resolutions (length of pixel edge).

Table 1 summarizes the comparisons of the GPW and DMSP OLS models of population density to both the spatially aggregated and smoothed standards. The standards are the 1-km^2 measure of population density derived from the 1990 U.S. census block group polygons and the 5 by 5 mean filter smoothed version of that dataset. The DMSP model estimates outperformed the global demography project estimates in almost every category at all scales, including the scale of the global demography project data.

A visualization of the spatial distribution and magnitude of the errors of these models does shed some light on the limitations of these representations of population density. Both the smoothed and aggregated standards to which the GPW and DMSP OLS estimates were compared did not include empirical information about where people work, shop, travel, etc. Plate 1 is an image of the errors of the DMSP OLS nighttime image-based model compared to the unsmoothed population density standard in the Los Angeles area. The errors manifest from spatial mis-allocations of population density. Overestimates (blue) happen in well-lit areas that have low population densities according to the census data. Underestimates (red) occur in high population density central business districts.

Many of the errors are clearly interpretable. For example, the population density of urban centers is typically

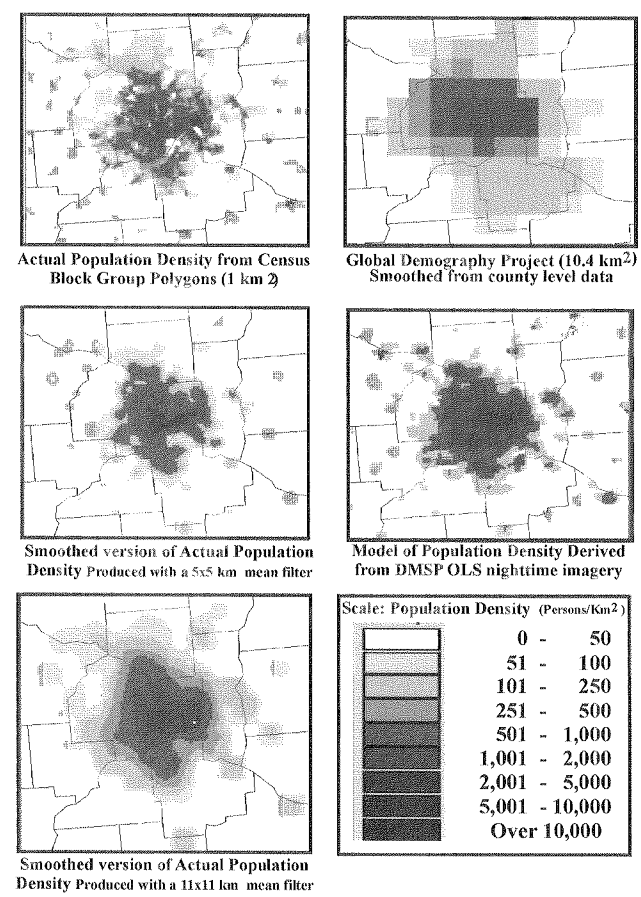

Figure 3. Representation of the population density of Minneapolis-St.Paul from various versions of datasets described in this paper.

TABLE 1. COMPARISONS OF GPW AND DMSP OLS MODELS TO STANDARDS FOR ALL U.S. URBAN AREAS

GDP = Global Demography Project	Standard: Actual Population Density (No Smoothing)					
DMSP = Model From Nighttime Imagery	Pixel Size: 1.0 km		Pixel Size: 5.0 km		Pixel Size: 10.4 km	
Figures from Linear Regression	GDP vs. Standard	DMSP vs. Standard	GDP vs. Standard	DMSP vs. Standard	GDP vs. Standard	DMSP vs. Standard
R^2	0.05	0.318	0.13	0.414	0.168	0.52
Mean Error	120.3	−60.7	119.1	−0.75	92	−106.6
GDP = Global Demography Project	Standard: Smoothed Pop. Den. (5 × 5 km Mean Filter)					
DMSP = Model From Nighttime Imagery	Pixel Size: 1.0 Km		Pixel Size: 5.0 km		Pixel Size: 10.4 km	
Figures from Linear Regression	GDP vs. Standard	DMSP vs. Standard	GDP vs. Standard	DMSP vs. Standard	GDP vs. Standard	DMSP vs. Standard
R^2	0.11	0.49	0.43	0.49	0.14	0.54
Mean Error	116.6	−64.3	52.2	−85.7	33.46	−92.5

underestimated. These areas tend to be well lit in the image; however, the linear proportion of population allocated to these areas based on light intensity is insufficient to account for the true population density. These errors could be mitigated with the use of a non-linear allocation of population density (Sutton, 1997); however, this would increase the existing errors of overestimation. The two large airports in the cluster are significant overestimates. The area around the Los Angeles International Airport (LAX) shows up as a large error of overestimation of population density. LAX is one of the largest employers in the city and tens of thousands of people fly in, out, or through that airport on a daily basis, yet the area has a relatively low population density according to the residence-based data derived from the census. The same is true of Ontario International Airport.

The interpretable and spatially non-random nature of these errors suggests that the smoothing and aggregation manipulations of the residence-based standard of population density are inadequate measures of the ambient population density that these models are trying to capture. Theoretically, a better standard of ambient population density could be derived from the temporally averaged spatial behavior of all persons present (residents and visitors) in the Los Angeles area for that given time. The financial and practical obstacles to obtaining such a dataset are presently insurmountable; however, an improved standard of ambient population density was derived from a combination of the 1990 residence-based measure of population density and a 1990 employment-based measure derived from PUMS data.

Spatial aggregation and mean filtering are somewhat *ad hoc* attempts at producing an ambient population density standard from a residence-based measure of population density. They are isotropic in nature, and knowledge of human spatial behavior tells us that ambient population density cannot be derived from an isotropic manipulation of residence-based population density. The final standard of ambient population density was derived by simply averaging a residence-based measure and an employment-based measure of population density for the Los Angeles area. This manipulation incorporates additional empirical data and consequently captures some of the anisotropic nature of the difference between residence-based population density and ambient population density.

The estimate of ambient population density derived from the DMSP OLS nighttime imagery was the only estimate whose correlation with the standard derived from both residence- and employment-based data was higher than its correlation with either the residence- or employment based measure alone. Plate 2 summarizes the comparisons of the DMSP OLS derived model and the residence- and employment-based standards in the Los Angeles area. It is interesting to note that the spatial pattern of error in Plate 2 is more random than the error derived from just the residence-based measure shown in Plate 1.

An interesting statistical question to ask is: '*What is the expected value of the correlation between dataset A and the mean of datasets B & C?*' A naïve guess might be that Correlation(A, ½ B + ½ C) would simply be the mean of Correlation(A, B) and Correlation(A, C). This is not the case. The actual expected value is given by

$$\begin{aligned} \mathrm{COR}(A, \tfrac{1}{2}B + \tfrac{1}{2}C) = &(\sigma_b^2/\sigma_{b+c}^2)^*\mathrm{COR}(A,B) \\ &+ (\sigma_c^2/\sigma_{b+c}^2)^*\ \mathrm{COR}(A,C) \\ &+ 2\mathrm{Cov}(A,B)\mathrm{Cov}(A,C)/(\sigma_a^2/\sigma_{b+c}^2). \end{aligned} \quad (2)$$

For a derivation of Equation 2, see Appendix A. A correlation analysis of the DMSP OLS, GPW, and LandScan models with the residence- and employment-based standards of population density in the Los Angeles metro area was conducted to assess the accuracy of these three representations of ambient population density (Figure 4 and Tables 2 and 3). The GPW and DMSP OLS models had the strongest correlations with the ambient standard of population density derived from the mean of the residence- and employment-based measures of population density. GPW had the highest correlation with the residence-based measure and the DMSP OLS model had the highest correlation with the employment-based measure. The DMSP OLS model was the only model whose correlation with the ambient standard was higher than its correlation with either the residence- or employment-based standard alone. The DMSP OLS model also had the largest difference between the expected value of its correlation with the ambient standard (as calculated from Equation 2) and its observed value. Although it should be noted that each of the models (GPW, DMSP OLS, and LandScan) had correlations with the ambient standard that were significantly higher than their expected value.

Discussion

Most existing measures of population density are derived from the arbitrary spatial units of a national census and represent residential population density. The temporal quality of these measures are twofold in the sense that they probably best measure a nighttime population density, and the repeat cycle of the measurement is determined by the frequency of the census. While this is a legitimate and useful measure of population density, there are several drawbacks to this measure that the models described here can mitigate. Some of the advantages of the GPW-, LandScan-, and DMSP OLS-based measures of population density are (1) uniform spatial units that meet needs for studies of the human dimensions of global change, (2) a temporally averaged measure of population density based on human mobility, and (3) variants of these models that can provide global coverage on a more frequent repeat cycle (~1 year).

The three representations of population density described here—GPW, LandScan, and the DMSP OLS-based

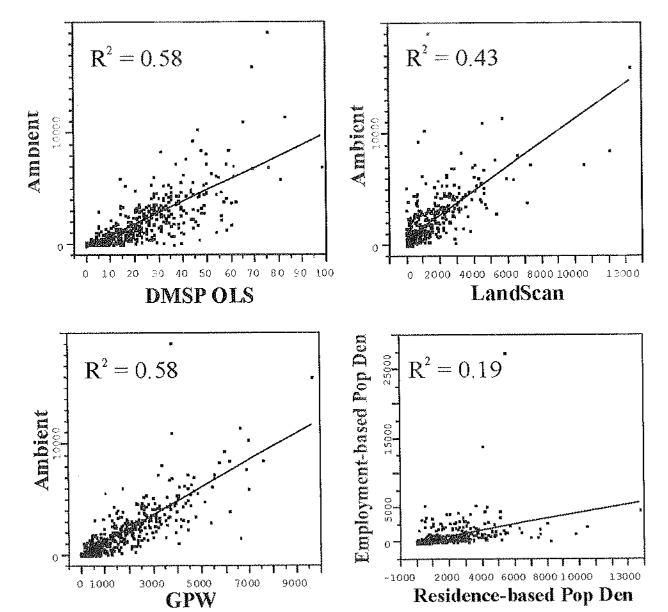

Figure 4. Scatterplots of correlation analysis.

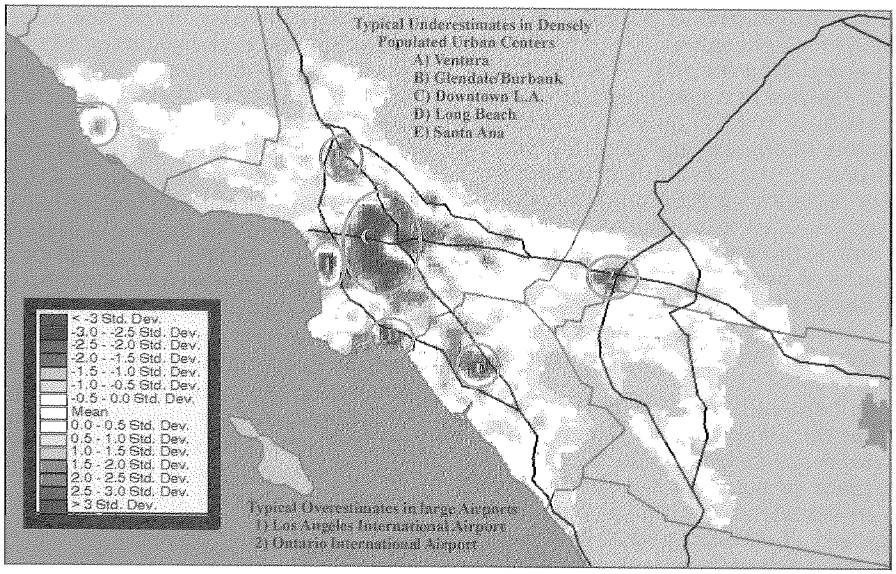

Plate 1. Image of the error (actual minus predicted) for the Los Angeles urban cluster.

model—have different strengths. GPW and LandScan both cover rural and urban areas whereas the DMSP OLS model is only developed for urban areas. The spatial resolutions of the DMSP OLS and LandScan are finer that that of GPW. GPW is in essence a re-representation of existing census data, whereas LandScan is essentially a manipulation of GPW that increases the spatial variability and resolution of the data and attempts to capture ambient as opposed to residence-based population density. On the other hand, the DMSP OLS model uses only one variable and can be imple-

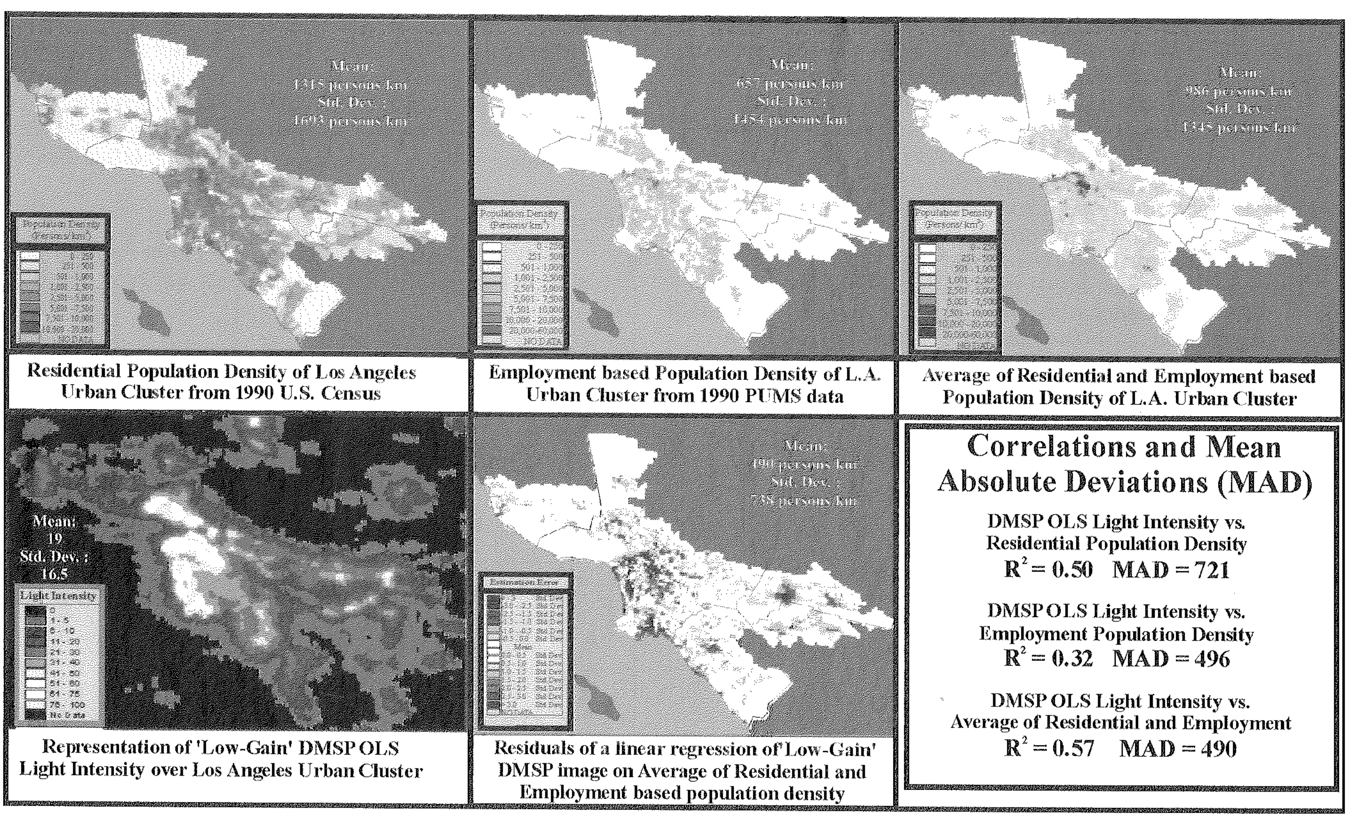

Plate 2. Comparisons of DMSP OLS based image of nighttime light intensity over Los Angeles to residential- and employment-based measures of population density.

TABLE 2. CORRELATION MATRIX OF MODELS AND STANDARDS

	Correlation (R) Matrix of Models and Standards					
Variable	DMSP OLS	LandScan	GPW	Residence	Employment	Ambient
DMSP OLS	1.00	0.58	0.76	0.72	0.57	0.76
LandScan	0.58	1.00	0.71	0.80	0.30	0.66
GPW	0.76	0.71	1.00	0.86	0.42	0.76
Residence	0.72	0.80	0.86	1.00	0.43	0.85
Employment	0.57	0.30	0.42	0.43	1.00	0.84
Ambient	0.76	0.66	0.76	0.85	0.84	1.00

TABLE 3. EXPECTED AND OBSERVED VALUES OF R^2 AND PARAMETERS NECESSARY FOR CALCULATION

Expected Value			Observed Value		
$R^2_{\text{DMSP OLS, Ambient}}$		0.29	$R^2_{\text{DMSP OLS, Ambient}}$		0.58
$R^2_{\text{LandScan, Ambient}}$		0.26	$R^2_{\text{LandScan, Ambient}}$		0.43
$R^2_{\text{GPW, Ambient}}$		0.33	$R^2_{\text{GPW, Ambient}}$		0.58
$\sigma^2_{\text{DMSP OLS}}$	1273839	Cov(DMSP OLS, Residence)	0.00000044	$R^2_{\text{DMSP OLS, Residence}}$	0.51
$\sigma^2_{\text{LandScan}}$	2370984	Cov(DMSP OLS, Employment)	0.00000030	$R^2_{\text{DMSP OLS, Employment}}$	0.32
σ^2_{GPW}	2217829	Cov(LandScan, Residence)	0.00000034	$R^2_{\text{LandScan, Residence}}$	0.64
$\sigma^2_{\text{Residence}}$	2890000	Cov(LandScan, Employment)	0.00000025	$R^2_{\text{LandScan, Employment}}$	0.09
$\sigma^2_{\text{Employment}}$	2685701	Cov(GPW, Residence)	0.00000037	$R^2_{\text{GPW, Residence}}$	0.74
$\sigma^2_{\text{Ambient}}$	1993744	Cov(GPW, Employment)	0.00000027	$R^2_{\text{GPW, Employment}}$	0.18

mented as an independent estimate of population and population density.

Recognition of the relative strengths and weaknesses of these models allows for their continued independent improvement and for potential synergy among them. Population density data is one of the fundamental datasets to which additional social, demographic, and economic data can be attached. Environmental data in uniform global coverage are increasing in quality, quantity, and usability; however; social, economic, and demographic data do not always lend themselves to immediate integration with these environmental datasets. Despite the great deal of important socio-economic and demographic data that are being collected, the incommensurate spatial units, reporting methods, and spatial and temporal scales of the information leave many analyses un-performable. In fact, several organizations and institutions have determined that socio-economic data at spatial and temporal scales commensurate with existing environmental data are the most significant need for investigating the human dimensions of global change (Clark and Rhind, 1992). Synergy between GPW-, LandScan-, and DMSP OLS-based models can serve as a vehicle for addressing this significant data gap.

GPW should remain as a gridded representation of demographic data derived from the "clean," "best estimate" administrative boundaries of official, nationally reported figures at the finest spatial resolution available. GPW could serve as part of the data framework for the development of a United Nations Geographic Database that is currently being discussed (United Nations, 2000). Ideally, this effort will focus on filling the gap in the areas of social, economic, and demographic information, particularly in developing countries.

LandScan is fundamentally derived from digital satellite imagery and cartographic products that are of finer spatial and temporal resolution than are traditional census data. This cartographic and remotely sensed information is cheaper to obtain and easier to process than are traditional census data. Theory needed to convert these data into estimates of additional social, economic, behavioral, and demographic attributes other than population density is progressing rapidly.

The DMSP OLS imagery is an excellent case in point. Global DMSP OLS imagery is being used to produce fine spatial resolution maps of human settlements (Elvidge *et al.*, 1997a), map urban extent nationally (Imhoff et al. 1997), estimate urban populations nationally (Sutton *et al.*, 2001), and serve as a proxy measure of economic activity, energy consumption, and CO_2 emissions (Elvidge *et al.*, 1997; Doll *et al.*, 2000). This research demonstrates the potential of the DMSP OLS imagery as a proxy measure of ambient population density. The potential of these proxy measures of social, economic, and demographic information is only just being exploited.

Conclusion

There are many potential benefits provided by a greater understanding of the relationship between remotely sensed imagery and ground-based measures of socio-economic and demographic information. Many countries of the world lack the financial and/or institutional resources to conduct useful censuses. Models derived from readily available satellite imagery that have been validated in parts of the world where good ground based information is available could serve as reasonable proxy measures in countries that lack such information. In addition, if a country has some limited resources with which to conduct an incomplete census of its population, existing imagery for that country could be used to help design statistical sampling strategies for a limited census. These sampling strategies could be designed to maximize the effectiveness and accuracy of proxy measures derived from satellite imagery. These methods could also be used to provide inter-censal estimates. This symbiosis allows for and benefits from independence of ground and imagery based measures.

These proxy measures of socio-economic and demographic information provided by remotely sensed imagery are easily incorporated into a myriad of environmental analyses such as global warming, deforestation, loss of biodiversity, etc. The finer spatial and temporal resolution of these proxy measures of population will be useful in many studies for planning, mitigation, and response to natural and anthropogenic disasters. If these kinds of models are developed and maintained over time, the time series data could be used for validation of urban growth models, planning for habitat preservation, and informing "smart growth" initiatives. Establishing models that estimate socio-econo-demographic information from remotely sensed imagery also has the potential to increase cooperation among nations and between those cooperating nations and the United Nations.

The GPW, LandScan, and DMSP OLS models of ambient population density described here all make foundational contributions to future efforts at filling the gap in social, economic, and demographic information for parts of the world where such data are unavailable. GPW led the way by providing global coverage of residence-based population density information in a format that is easily incorporated into grid- or raster-based environmental models. LandScan switched gears by moving to a representation of ambient population density by incorporating additional finer spatial resolution digital information such as roads, topography, and nighttime satellite imagery to allocate population density to that finer spatial resolution. LandScan is probably the best global representation of rural and urban population density available today. The model of ambient population density derived from the DMSP OLS imagery is perhaps the best representation of ambient population density in urban areas at the spatial resolution of LandScan. The DMSP OLS imagery has also been used to estimate other phenomena such as GDP, CO_2 emissions, and urban extent. In the future we believe that synergy between these three models will allow for development of methods to independently estimate other hard to obtain social, economic, and demographic information.

Acknowledgments
This research was supported in part by the department of Geography at the University of Denver and an NSF grant (#CMS-9817761, An integrated Modeling Environment for Urban Change Research). All support is gratefully acknowledged. Constructive criticism from various anonymous peer reviewers is thankfully acknowledged. We would also like to thank Michael Goodchild and Waldo Tobler for helpful comments on the manuscript in its earlier stages.

References
Clark, J., and D. Rhind, 1992. *Population Data and Global Environmental Changes*, International Social Science Council, Programme on Human Dimensions of Global Environmental Change, UNESCO, Paris, France.

Deichmann, U., 1996. *A Review of Spatial Population Database Design and Modeling*, National Center for Geographic Information and Analysis, University of California at Santa Barbara, Santa Barbara, California, 58 p.

Dobson, J.E., E.A. Bright, P.R. Coleman, R.C. Durfee, and B.A. Worley, 2000. LandScan: A global population database for estimating populations at risk, *Photogrammetric Engineering & Remote Sensing*, 66(7):849–857.

Doll, C.N.H., J.-P. Muller, and C.D. Elbidge, 2000. Night-time imagery as a tool for global mapping of socio-economic parameters and greenhouse gas emisssions, *Ambio*, 29(3):159–164.

Elvidge, C., K. Baugh, E. Kihn, H. Kroehl, E. Davis, and C. Davis, 1997a. Relationship between satellite observed visible-near infrared emissions, population, economic activity, and electric power consumption, *International Journal of Remote Sensing*, 18:1373–1379.

Elvidge, C.D., K.E. Baugh, V.H. Hobson, E.A. Kihn, H.W. Kroehl, E.R. Davis, and C. Davis, 1997b. Satellite inventory of human settlements using nocturnal radiation emissions: A contribution for the global toolchest, *Global Change Biology*, 3:387–395.

Elvidge, C.D., K.E. Baugh, J.B. Dietz, T. Bland, P.C. Sutton, and H.W. Kroehl 1998. Radiance Calibration of dmsp-ols low-light imaging data of human settlements, *Remote Sensing of Environment*, 68:77–88.

Haub, C., and D. Cornelius, 1998. *World Population Data Sheet*, Population Reference Bureau, Washington, D.C.

Holloway, L.J., Jr., 1958. Smoothing and filtering of time series and space fields, *Advances in Geophysics*, 4, 351-389

Imhoff, M.L., W.T. Lawrence, C.D. Elvidge, T. Paul, E. Levine, and M.V. Privalsky, 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(1):105–117.

Landford, M., and D.J. Unwin, 1994. Generating and mapping population density surfaces within a geographical information system, *The Cartographic Journal*, 31(June):21–25.

Long, L.H., and C.G. Boertlein, 1976. *The Geographical Mobility of Americans: An International Comparison*, Bureau of the Census, Washington, D.C., 65 p.

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

Sutton, P.C., D. Roberts, C.D. Elvidge, and K. Baugh, 2001. Census fromhHeaven: An estimate of the global population using nighttime satellite imagery, *International Journal of Remote Sensing*, 22(16):3061–3076.

Tobler, W., 1979. Smooth pycnophylactic interpolation for geographic regions, *Journal of the American Statistical Association*, 74(367):519–536.

Tobler, W.R., U. Deichmann, J. Gottsegen, and K. Malloy, 1995. *The Global Demography Project, National Center for Geographic Information and Analysis*, University of California at Santa Barbara, Santa Barbara, California, 75 p.

———, 1997. World population in a grid of spherical quadrilaterals, *International Journal of Population Geography*, 3:203– 225.

United Nations, 2000. *Cartography and Geographic Information Science*, United Nations, New York, N.Y., 53 p.

(Received 23 October 2001; accepted 23 May 2002; revised 12 July 2002)

Appendix A—Derivation of Equation 2

$$\mathbf{COR(A,\ {}^1\!/_2\,B + {}^1\!/_2\,C) = (\sigma_B{}^2/\sigma_{B+C}{}^2)*COR(A,B) + (\sigma_C{}^2/\sigma_{B+C}{}^2)*COR(A,C) + 2Cov(A,B)Cov(A,C)/(\sigma_A{}^2*\sigma_{B+C}{}^2)}$$

$$\mathbf{COR(A,B) = R^2{}_{A,B}}$$

$$\mathrm{COR\ (A,\ {}^1\!/_2\,B + {}^1\!/_2\,C) = [Cov(A,\ {}^1\!/_2\,B + {}^1\!/_2\,C)^2/(\sigma^2{}_A)(\sigma^2{}_{{}^1\!/_2B\ +\ {}^1\!/_2C})}\ \textbf{(by definition)}$$

$$= \frac{{}^1\!/_4[\mathrm{Cov(A,B)}]^2 + {}^1\!/_4[\mathrm{Cov(A,C)}]^2 + {}^1\!/_2[\mathrm{Cov(A,B)Cov(A,C)}]}{(\sigma^2{}_A)[1/4\ \sigma^2_B + 1/4\ \sigma^2{}_C + {}^1\!/_2\ \mathrm{Cov(B,C)}]}$$

$$= \frac{[\mathrm{Cov(A,B)})]^2 + [\mathrm{Cov(A,C)}]^2 + 2[\mathrm{Cov(A,B)Cov(A,C)}]}{\sigma^2{}_A[\sigma^2_B + \sigma^2{}_C + 2\mathrm{Cov(B,C)}]}$$

$$= \frac{[\mathrm{Cov(A,B)})]^2 + [\mathrm{Cov(A,C)}]^2 + 2[\mathrm{Cov(A,B)Cov(A,C)}]}{\sigma^2{}_A\sigma^2{}_{B+C}}$$

$$\mathbf{COR(A,\ {}^1\!/_2\,B + {}^1\!/_2\,C) = (\sigma_B{}^2/\sigma_{B+C}{}^2)*R^2{}_{A,B} + (\sigma_C{}^2/\sigma_{B+C}{}^2)* + R^2{}_{A,C} + 2Cov(A,B)Cov(A,C)/(\sigma_A{}^2*\sigma_{B+C}{}^2)}$$

To Calculate COR(A, ½ B + ½ C), the following must be known:

1) $\rho_{A,B}$ or $\rho^2{}_{A,B} = R^2{}_{A,B} = \mathrm{COR(A,B)}$
2) $\rho_{A,C}$ or $\rho^2{}_{A,C} = R^2{}_{A,C} = \mathrm{COR(A,C)}$
3) $\mathrm{Cov(A,B)} = \rho_{A,B}{*}\sigma_A{*}\sigma_B$
4) $\mathrm{Cov(A,C)} = \rho_{A,C}{*}\ \sigma_A{*}\sigma_C$
5) $\sigma^2{}_{B+C} = 4{*}\ \sigma^2{}_{B+C/2}$
6) $\sigma^2{}_A$
7) $\sigma^2{}_B$
8) $\sigma^2{}_C$

Photogrammetric Engineering and Remote Sensing (PE&RS)

Submitting a Manuscript for Peer Review ● Instructions for Authors

Authors submitting a new manuscript for peer review should follow these instructions.

Introduction: The American Society for Photogrammetry and Remote Sensing (ASPRS) seeks to publish in **Photogrammetric Engineering and Remote Sensing (PE&RS)** theoretical and applied papers that address topics in photogrammetry, remote sensing, geographic information systems (GIS), the Global Positioning System (GPS) and/or other geospatial information technologies. Contributions of manuscripts that deal with technical advancements in instrumentation, novel or improved modes of analysis, or innovative applications of these technologies in natural and cultural resources assessment, environmental modeling, or the Earth sciences (atmosphere, hydrosphere, lithosphere, biosphere, or geosphere) are especially encouraged.

Review Procedures: Manuscripts are peer reviewed and refereed by a panel of experts selected by the Editor (or Guest Editor - see special note below). Manuscripts are peer reviewed using a blind review procedure. The identities and affiliations of authors are not provided to reviewers, nor are reviewers' names disclosed to authors. The **PE&RS** goal is to complete reviews within 90 days of receipt of a manuscript by the Editor. Manuscripts accepted for publication will be returned to the author(s) for final editing before being placed in the queue for publication. Manuscripts not accepted will either be (1) rejected or (2) returned to the author(s) for revision and subsequent reconsideration by the review panel. Authors who do not revise and return a "to-be-reconsidered" manuscript within 90 days from receipt of reviews may have their manuscript withdrawn from the review process.

English Language: Authors whose first language is not English should have their manuscripts reviewed by an English-speaking colleague or editor to refine use of the English language (vocabulary, grammar, syntax). At the discretion of the Editor, manuscripts may be returned for English language corrections before they are sent for review.

Format Requirements: Manuscripts submitted for peer review must be prepared as outlined below. Papers that do not conform to the requirements described below will be returned for format revisions before they are sent for review. For a sample of a correctly prepared manuscript, please see peer reviewed articles published in any recent issue of **PE&RS**.

1 **Typing:** Manuscripts must be double-spaced (only 3 lines to the vertical inch) on one side of 8½- by 11-inch or A4 International (210- by 297-mm) white paper, with 30-mm (1¼-inch) margins all around. For review purposes every part of the manuscript must be double-spaced, including title page, abstract, text, footnotes references, appendices, and figure captions. Single-spaced manuscripts will be returned.

2 **Number of Copies**: Four (4) hard copies of each manuscript are required. Each copy should include all illustrations, tables and line drawings. For line drawings, an original copy, plus three (3) xerographic (or equivalent) copies is acceptable. Color photocopies of color images are required for review.

3 **Paper Length**: Authors are encouraged to be concise; lengthy papers are discouraged. Published papers are generally limited to eight (8) journal pages. A 30-page manuscript (including tables and figures), when typed as indicated above, equals about eight journal pages. Authors of published papers will be charged $100/page for each page exceeding eight journal pages. These page charges must be paid before publication.

4 **Title Page**: Authors should strive for titles no longer than eight to ten words. The title page shall include a short title and a one-sentence description of the paper's content to accompany the title in the **PE&RS** Table of Contents. To facilitate the blind review process, **authors' names, affiliations, and addresses should be listed only on the cover letter, not on the title page**. Authors should indicate both their current affiliation and, if different, their affiliation at the time the research was performed.

5 **Abstract**: All manuscripts submitted for peer review (except "Briefs," i.e., manuscripts of fewer than 5 typewritten pages) must include an abstract of 150 words or less. The abstract should be complete, informative, succinct and understandable without reference to the text.

6 **Figures and Tables**: All figures and tables must be cited in the text. Authors should note that figures and tables will usually be reduced in size by the printer to optimize use of space, and should design accordingly. For purposes of peer review, figures and tables should not be embedded in the text, but should be appended at the end of the text. Captions must be double-spaced on a single page preceding the tables and illustrations. If the manuscript contains copyrighted imagery, a copyright statement must be included in the caption (e.g. ©SPOT Image, Copyright [year] CNES).

7 **Color Illustrations**: Authors should use black and white illustrations whenever possible. Authors who include color illustrations will be charged for the cost of color reproduction. These costs must be paid before an article is published. Additional information on color reproduction costs will be forwarded to authors upon request or when a manuscript is accepted for publication. Note that in some cases, authors may wish to publish black-and-white versions of color illustrations in the printed journal, and post electronic versions (in full color) of the same illustrations on the ASPRS web site (see #10 Electronic Journal below). If this is intended, it must be made clear in the manuscript. Color illustrations must still be provided for review purposes.

8 **Metric System**: The metric system (SI Units) will be employed throughout a manuscript except in cases where the English System has special merit stemming from accepted conventional usage (e.g., 9- by 9-inch photograph, 6-inch focal length). Authors should refer to "Usage of the International System of Units," **Photogrammetric Engineering & Remote Sensing**, 1978, 44 (7): 923-938.

9 **Equations**: Authors should include only those equations required by a typical reader to understand the technical components in the manuscript. Lengthy derivations, or mathematics that have already been published in **PE&RS**, or elsewhere, should not be included in the text file (but, see **Electronic Journal** below). Authors should express equations as simply as possible. Whenever possible, authors are encouraged to use the **Insert** and **Symbol** capabilities of *Microsoft Word* to build simple equations. If that is not possible, authors are encouraged to use *Microsoft Equation 3.0* (or later), or *Microsoft Equation Editor* for developing equations. The author corresponding with the Editor must indicate in the letter of transmittal which software was used to create the equations. Equations must be numbered, but unlike tables, figures, color plates, and line drawings, should be embedded in the text file.

10 **Electronic Journal**: The ASPRS Journal Policy Committee discourages lengthy appendices, complex mathematical formulations, and software programs. These will ordinarily not be published in the hardcopy version of **PE&RS**. However, these materials may be made available on the ASPRS web site (www.asprs.org). Some authors may also wish to publish color illustrations (e.g., color versions of illustrations printed in black and white in the hard copy journal) on the web site. Authors wishing to have supplemental material or color illustrations posted on the website after their paper is published should submit this material along with their manuscript. All supplemental material must be clearly labeled as *supplemental material*.

11 **References**: A complete and accurate reference list is essential. Only works cited in the text should be included. Cite references to published literature in the text by authors' last names and date, as for example, Jones (1979), Jones and Smith (1979) or (Jones, 1979; Jones and Smith 1979), depending on sentence construction. If there are more than two authors, they should be cited as Jones *et al.* (1979) or (Jones *et al.*, 1979). Personal communications and unpublished data or reports should not be included in the reference list but should be shown parenthetically in the text (Jones, unpublished data, 1979). Format for references in the reference list will be as follows:

Books:
Falkner, E., 1995. *Aerial Mapping: Methods and Applications*, Lewis Publishers, Boca Raton, Florida, 322 p.

Articles (or Chapters) in a Book:
Webb, H., 1991. Creation of digital terrain models using analytical photogrammetry and their use in civil engineering, *Terrain Modelling in Surveying and Civil Engineering* (G. Petrie and T.J.M. Kennie, editors), McGraw-Hill, Inc., New York, N.Y., pp. 73-84.

Journal Articles:
Meyer, M.P., 1982. Place of small-format aerial photography in resource surveys, *Journal of Forestry*, 80(1):15-17.

Proceedings (printed):
Davidson, J.M., D.M. Rizzo, M. Garbelotto, S. Tjosvold, and G.W. Slaughter, 2002. *Phytophthora ramorum* and sudden oak death in California: II. Transmission and survival, *Proceedings of the Fifth Symposium on Oak Woodlands: Oaks in California's Changing Landscape*, 23-25 October 2001, San Diego, California (USDA Forest Service, General Technical Report PSW-GTR-184, Pacific Southwest Forest and Range Experiment Station, Berkeley, California), pp. 741-749.

Proceedings (CD-ROM):
Cook, J.D., and L.D. Ferdinand, 2001. Geometric fidelity of Ikonos imagery, *Proceedings of the ASPRS 2001 Annual Convention*, 23-27 April, St. Louis, Missouri (American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland), unpaginated CD-ROM.

Thesis and Dissertations:
Yang, W., 1997. *Effects of Spatial Resolution and Landscape Structure on Land Cover Characterization*, Ph.D. dissertation, University of Nebraska-Lincoln, Lincoln, Nebraska, 336 p.

Website References:
Diaz, H.F., 1997. Precipitation trends and water consumption in the southwestern United States, USGS Web Conference, URL: http://geochange.er.usgs.gov/sw/changes/natural/diaz/, U.S. Geological Survey, Reston, Virginia (last date accessed: 15 May 2002).

12 **Acknowledgments**: In keeping with the process of blind reviews, authors are asked not to include acknowledgments in manuscripts submitted for peer review. An acknowledgment may reveal a considerable amount of information for reviewers that is not necessary or desirable for their evaluation of the manuscript. After a manuscript is accepted for publication, the lead author will be encouraged to insert appropriate acknowledgments.

Manuscript Submission:: Manuscripts submitted for peer review should be mailed to Dr. James W. Merchant, Editor, Photogrammetric Engineering and Remote Sensing
113 Nebraska Hall, 901 North 17th Street, University of Nebraska-Lincoln
Lincoln, NE 68588-0517 USA
E-mail: jmerchant1@unl.edu, Tel.: 402/472-7531

Note: Authors should NOT MAIL MANUSCRIPTS TO ASPRS HEADQUARTERS. This will cause the review to be delayed.

Instructions last updated July 17, 2002