4 Intraurban Population Estimation Using Remotely Sensed Imagery

Perry J. Hardin, Department of Geography, Brigham Young University, 690 SWKT, Provo, UT 84602, USA

Mark W. Jackson, Department of Geography, Brigham Young University, 690 SWKT, Provo, UT 84602, USA

J. Matthew Shumway, Department of Geography, Brigham Young University, 690 SWKT, Provo, UT 84602, USA

Of the Earth's 6.5 billion human inhabitants, nearly three billion live in urban settlements (UNCHS, 2001). Natural increase, land tenure practices, political policy, environmental degradation, and the dynamics of regional / global economics are largely responsible for the ongoing population shift from rural agrarian regions to cities. This increased urbanization is not just a developing country phenomenon. Urban areas of North America in 1900 were home to only 50% of the continent's population. In 2000, the percentage of North American urban inhabitants rose to 75%.

Given the importance of urban regions as human habitat, there is an established need for accurate intraurban population counts to support decision making. In comparison to population estimates or projections, an exhaustive per-dwelling enumeration acquired through fieldwork is the accepted gold standard for counting people and determining their sociodemographic characteristics. A census is a complex undertaking; it requires significant human, technological, and fiscal resources to plan and execute. Because of high cost, industrialized nations conduct enumerations only periodically. The American decennial census mandated by the U.S. constitution is an example. The United Kingdom also conducts census surveys every ten years. The Australian Bureau of Statistics conducts a census every half decade. In contrast, developing nations experience almost insurmountable obstacles to obtain accurate and regular enumerations of their national population. These include vast rural areas, nomadic populations, functional sovereignty limitations, cultural mistrust, and a lack of technical and financial wherewithal.

The Decennial Census of the US is intended as a temporal snapshot; a depiction of the national population on April 1 of the census year. In areas of rapid urban growth, the population counts recorded in a decennial census become progressively less representative as the decade progresses. Recognizing this, intercensal population estimations are commonly required. Additionally, small area population estimates provide a key source of data for local planning agencies and businesses. Often the data are required at a geographic level smaller than what is easily found in census data. In these cases, population estimates are the most cost-efficient way to generate the required small area data.

In this chapter we will first briefly review the traditional methods for population estimation. Three broad methodologies for estimating intraurban population totals and densities using overhead imagery will then be discussed. Our focus is on the developed urban world rather than developing or rural areas. Remote sensing may provide a vital role for population estimation in developing countries with significant rural regions, but it is not our expertise.

Once the three broad methodologies have been appraised, a short case study will be presented. In this case study, we use rudimentary image processing techniques to estimate the population of the Wasatch Front urban corridor in Utah, U.S. After the case study, some concluding comments are then offered about future research directions.

4.1 Traditional Approaches to Population Estimation

Population estimates should not be confused with population projections. Although data and methods may differ, the primary difference is one of time period. Population *estimates* are used for the present and the past, whereas population *projections* are used to guesstimate future population size. In this chapter, our focus is on population estimates.

Estimating population of small areas at various scales of space and time is a difficult demographic task. However, because small area population estimates are often necessary for local planning departments and businesses (billions of dollars in federal funds are allocated to states and local entities based on the estimates), there have been several estimation methodologies developed. Regardless of the method, four preliminary factors must be considered; 1) the purpose of the estimation, 2) the spatial scale of the estimation, 3) the target temporal period for the estimation, and 4) data availability. Once those factors have been considered, three other issues must likewise be addressed; 1) the collection of any necessary data, 2) the selection of correct statistical methods, and 3) the method for judging the goodness of the estimates.

The collection of necessary data is framed by the purpose and geographic scale of the required estimate. At smaller geographic scales certain administrative records are aggregated at a county or metropolitan scale and may not be available at sub-county levels (e.g. tax returns). Deciding the scale and purpose of the estimate often determines what data can be used, which in-turn constricts the choice of appropriate methods. However, almost all traditional population estimation techniques use various types of administrative records that are correlated with population change. Predictors derived from these records are called *symptomatic variables* (Plane and Rogerson 1994). Ideally, symptomatic variables should be updated regularly. They should also temporally co-vary with population change in a predictable fashion. Exemplar symptomatic variables include residential building permits, utility connections, school enrollments, tax returns, and Medicare enrollments. A second important data source is vital records – particularly birth and death certificates. These data are used as major inputs into the cohort-component method of population estimation described below.

The U.S. Census Bureau, in cooperation with state partners, is legally required to provide intercensal population estimates to support federal fund allocations. To comply, the Census Bureau has developed three principle methods:

1. *Ratio-correlation procedures*. As the name implies, ratio-correlation procedures use the ratio of symptomatic variable values for adjacent time periods as independent and dependent variables to estimate population. Changing ratios of symptomatic variables within a geographic region are assumed to be a function of the region's changing population ratio (Plane and Rogerson 1994). The ratio of a subregion's population to the larger region population for two time periods is regressed on similarly formed symptomatic variables (ratio of ratios). These models will generally use vital records (i.e., births

and deaths) and administrative records (e.g.,elementary school enrollment, vehicle registration, voter registration).

2. *Component-method II procedures*. All component methods are predicated on demographic accounting, i.e. population change = births - deaths + net migration. The greatest difficulty with this approach is correctly specifying the migration factor. The component-method II procedure utilizes registration data on births and deaths, and tries to estimate net migration using other administrative information. The U.S. Census Bureau splits the use of different administrative data genres based on age. For populations less than 65 years of age individual tax returns are used.[1] For populations older than 65 years of age, medicare enrollment is used. In this method, migration – based on different tax return addresses or medicare payment addresses – is estimated. Both sources also have information on household size. For entities without access to individual tax returns, school enrollment is often used and assumed to be indicative of migration in the total population – with adjustments being made for the historical differences between the school-age migration rate and the total population's rate of migration.

3. The *housing-unit method* is based on change in the housing stock of an area from the base date to the estimate date. Data on the housing stock and flow are generally derived from; 1) U.S. Bureau of the Census survey of building permits and demolitions, and 2) State Data Center survey of counties and cities issuing permits for residential buildings and demolitions. The housing unit method requires the specification (assumption) of vacancy rates and average household size. Once specified, housing unit count change between base and estimate dates is multiplied by the occupancy rate and average household size to estimate population change. Individuals in group quarters (prisons, college dormitories, nursing homes, and military barracks) are included in the total. As a refinement, separate estimates are constructed by housing structure type (e.g., single-family dwellings, 2-to-4 unit, 5+ units, mobile homes). This refinement permits different vacancy and household size factors to be more precisely tailored to the structure types within the housing stock.

---

1 Only the Census Bureau has access, from the IRS, to the individual tax returns. Other governmental or private businesses will use school enrollment data in lieu of individual tax returns.

Perhaps the most important aspect of population estimation is validation. Population estimates are typically based on assumptions of temporally stable relationships between population change (i.e., births, deaths, migration) and their symptomatic variables. The temporal stability of the symptomatic variables themselves is likewise assumed. These are generally safe assumptions, but the estimate will always contain error. The only way to assess the error is to do an actual count. This would of course obviate the need for the estimate in the first place.

4.2 Population Estimation Using Remote Sensing

There are four primary approaches to estimating population with remotely sensed data:

1. The use of allometric population growth models based on place size. Typically the area of cities, towns and villages is measured from small scale air photography or satellite imagery and submitted to a calibrated allometric model to estimate population. Central place theory and road connectivity are sometimes employed to improve accuracy. The allometric technique is very useful in developing countries where ground enumeration is impossible and a single population total for each city, village, or region is acceptable. It is less useful when population estimates are required for small enumeration districts such as US census tracts. This method is beyond the scope of this paper but Lo (2006) provides an excellent review.

2. The use of dwelling unit type as a surrogate for family size.[2] This technique requires an interpreter to identify, classify, and count dwelling units manually from large scale imagery. A simple model relates dwelling type to resident family size.

3. The use of landtype zones[3] as a surrogate for population density. Different landtype zones are identified on medium-scale imagery and

[2] With some loss of precision, we call this approach *dwelling identification*.

[3] Every student of remote sensing quickly learns the difference between land use and land cover. Nonetheless, to simplify phraseology by avoiding the repetition of the phrase "land use and/or land cover" throughout this chapter, the term

a model is employed that links landtype with population density. This approach has historically been used with medium scale air photography and satellite imagery.

4. Pixel based approaches that seek to model population or housing unit density directly as a function of spectral reflectance or spectral texture on a medium-scale satellite images.

4.2.1 Dwelling Identification

The process of estimating population density via dwelling identification is conceptually simple and requires the following general steps. A schema of dwelling unit types based on family size is initially itemized. For example, the schema may include designations such as duplexes, single family residential homes, and apartments. Using information acquired from census data, interviews, or rental agencies, average resident counts for each dwelling unit type in the schema are determined. Each dwelling unit in the study area is then placed into one of the *a priori* schema classes by its appearance on large scale photography. Total estimated population is the sum of the dwelling units of each type weighted by their corresponding average resident population.

The success of the procedure outlined above depends on the successful identification of various dwelling types from high-resolution imagery. This was established early by Green (1956, 1957) who postulated that the social structure of a city could be determined through the analysis of aerial photography. Green suggested that this identification of dwelling type is the first step to the use of air photography for demographic, sociological, and urban ecological applications. In this pioneering research, Green (1956) examined 17 residential neighborhoods in Birmingham, AL to ascertain whether stereo air photography (1:8,000 scale) facilitated the identification of urban dwelling structure type. Although the research focused on measuring; 1) the percentage of detached single-unit homes, and 2) the dwelling unit density per block, several other residential structural types were discriminated as part of the study (e.g., duplexes, multiunit). Green utilized the following characteristics in his photographic key to housing identification:

“landtype” will be used instead. Where differentiating between land cover and land use is important to the discussion, the two separate terms will be used.

1. Roof structure and form (i.e., gables, dormers, porches) including the existence and number of chimney stacks and rooftop plumbing fixtures.
2. The overall shape and size of the building.
3. The situation of the building i.e., “the location of the building with respect to the street, the building line, and other structures” (p. 143).
4. Vehicle accommodations, including carports, parking areas, garages, and driveways.
5. Pedestrian accommodations such as footpaths, sidewalks, and entryways.
6. The shape and size of yards, courts, etc.

Not surprisingly, Green found that the error rates in classification were dependent on structure type rather than uniformly distributed across all structure classes. Nearly all the problems involved multiunit residences.[4] Specifically, Green had trouble with universally distinguishing multiunit complexes from duplexes, and differentiating between duplexes and single-unit dwellings. Summarizing, the total study area housing unit count and multiunit structure count were slightly underestimated whereas the single-unit dwellings were overestimated. Overall however, “the results [showed] 1) that 99.8 percent of the 3,629 existing residential structures in the 228 city blocks observed were correctly identified as such, and 2) that 89 percent of these structures were correctly classified by categories of numbers of dwelling units.” For other related dwelling unit studies akin to Green (1956), see Hadfield (1963) and Binsell (1967).

Extending Green’s groundbreaking research, the objective of Lindgren’s (1971) study was to determine; 1) whether the same dwelling unit identification success reported in foregoing research could be obtained with medium-scale imagery (1:20,000), and 2) whether the use of color infrared (CIR) photography improved dwelling type identification success rates obtainable from natural color or panchromatic imagery. Lindgren’s operating assumption was that “in high-density areas, CIR imagery would allow for easier identification of urban signatures” (p. 374). Although originating with Binsell (1967), Lindgren’s final list of dwelling identification keys deviates little from Green (1956).

After developing the clues using three blocks of high-density housing in the metropolitan Boston area (i.e., East Boston, Chelsea, Charlestown), the

[4] The identification and treatment of multiunit structures is a reoccurring theme in population estimation with overhead imagery.

indicators were tested in 15 additional blocks containing 655 residential structures and 1744 dwellings. Lindgren's total residential structure count from the photography underestimated the actual total by only three. However, the ability to count infrastructure dwelling units was more difficult because of the prevalence of multiunit buildings within the study area – the interpretation underestimated the total number of dwelling units by 54. These summary figures obscure the seriousness of some interpretation problems – in both residential structure and dwelling unit counts, significant overestimates were offset by substantial underestimates. Overall, only 59% of the residential structures were assigned the correct number of dwelling units.

In concluding, Lindgren offered two observations. First, any personal familiarity of the study area enjoyed by the interpreter would dramatically increase chances for a successful outcome. For example, structures that Lindgren found in Charlestown with a particular roof-type were consistently mis-categorized. A single visit to Charlestown before the interpretation began would have prevented the mistake. Second, the high quality of the CIR transparencies used (i.e., their sharpness and contrast), in conjunction with the infrared distinction between built-up and nonbuilt-up urban areas more than counteracted any disadvantage of the small CIR image scale.

In coincident research, Collins and El-Beik (1971) used dwelling identification methods to estimate the population of the City of Leeds. The goal of the study was to determine whether population estimates made from air photography agreed with census estimates. Like researchers before them, the operational hypothesis was that dwelling type was strongly correlated with resident population count. Given earlier work by El-Beik (1967) demonstrating that housing types in Leeds could be identified classified from air photography, that hypothesis was reasonable.

The schema for the population estimation study required the discrimination of semidetached, terraced, and back-to-back dwelling types. Based on the interpretation of 1963 photographs of 1:10,000 scale, all the housing structures within the study area were classified into one of those three categories. Multiplication factors linking dwelling type to inhabitant number were derived from 1961 census enumeration maps and data. Only half of the enumeration districts were used to derive these factors whereas the other half was cloistered for validation purposes. For semidetached dwellings, it was found that 3.03 people lived in each house. The linear nature of semidetached and terraced dwellings in Leeds suggested a different tack for designating a multiplier. For these structures the multiplier was the total housefront linear feet per person with 4.62 and 4.27 ft / person being adopted for terraced and back-to-back houses respectively.

For validation, these factors were applied to the housing within the sequestered enumeration districts and population totals were calculated. When compared against 1961 census data, Collins and El-Beik found that the average error for enumeration districts dominated by terraced houses, back-to-back houses, and semidetached dwellings was +0.87%, +0.32%, and -6.4% respectively. The largest errors among individual enumeration districts were underestimates among semidetached homes with unexpectedly large families.

According to Collins and El-Beik, the accuracy of this approach depended primarily on two variables. The first was the ability of the photointerpreter to properly identify housing type. The second was how well the calibrated multiplier correctly represented target areas of the same dwelling category. The authors also observed that refinements in the method were possible. Multipliers could be adjusted according to structure age and proximity to schools. Social and economic variables could likewise be used to create a more sophisticated multiplier set.

Watkins (1984) focused his research on the problem of correctly counting the number of dwelling units in a multiunit structure. Error resulting from this prevalent problem was also investigated – Watkins observed that "no studies to date have explicitly investigated the nature of multiple dwelling unit counting errors with respect to the ways in which they relate to different structure types, nor have they considered the actual impact that multiple unit structures as a whole have on the accuracy of enumerations of all dwelling units within a residential area" (p. 1599).

Watkins subdivided multiunit structures into two groups; 1) those originally designed to house multiple families, and 2) structures originally built as single-family dwellings. Watkins observed that the diagnostic photographic elements needed to estimate the number of households were different in each group. The photographic key developed included not only guidelines for differentiating between residential and nonresidential but instructions for discriminating between converted single-family structures and archetypical apartment structures. Telltale features of apartment buildings included roof divisions, outside fire escapes and porches, entrance location and number, parking, and apparent socioeconomic level. Converted structures were distinguished by structure symmetry, quality and amount of vegetation, number of sidewalks from the structure to the roadway, walkways between sides of the same structure, site context, and the apparent method of property subdivision. Structure size, shape and height, as well as vents and chimneys were important to identifying both types.

After developing the photographic key employing these factors, Watkins conducted tests in Boulder, Colorado for three study areas. Dwelling count estimates made from 1:20,000 panchromatic imagery acquired in 1970 were compared to 1970 census block dwelling counts. Similarly, dwelling counts from the 1980 census were compared to estimates derived from 1979 1:6000 scale panchromatic photography. The photographic key was highly successful. For the 1970 data involving 695 buildings, errors in multiunit counts within the three study areas ranged from an underestimation of 1.61% to 0.37%. The error rates for 1980 (2545 buildings) were significantly higher, ranging from 1.64% to 4.91%. As hypothesized, error rates differed by multiunit structure type. Converted single dwelling units were overestimated by 8.45% whereas dwelling units within traditional apartment buildings were underestimated by 5.51%. Single units (unconverted) were underestimated by 4.80%.

Lo (1986) stands alone as the only researcher who has actually applied the dwelling unit identification method to an entire city. The goal of Lo's research was to estimate population in 93 traffic zones in Athens, Georgia from 1:20,000 aerial photography. Like preceding researchers, Lo used a simplified residential structure schema that included only a few structural types; 1) small single family structures, 2) large single family structures, and 3) multifamily structures. Estimated resident counts used in the population calculations were 3.0 and 4.0 for small and large family dwellings respectively, and 2.0 per dwelling unit within multifamily structures.

Comparison of the photointerpretation results with 1980 census data revealed an average population count underestimate of 1.7% per traffic zone with considerable variation from zone to zone. Counting errors were attributed to the following factors; 1) family sizes different from those assumed in the estimation process, 2) photointerpreter skill, 3) the number of multiunit structures, 4) the area of the traffic zone, and 5) the quality of the photographic source. Lo demonstrated that urban population estimation for an entire city was feasible, and resulting accuracy could be high. We consider Lo (1986) to represent the state-of-the-art in the dwelling unit identification / counting approach to intraurban population estimation.

Comments

Although the identification and counting of dwelling units showed significant promise for intraurban population estimation, there is little literature mentioning it after Lo (1986). Of all three methods for intraurban population estimation discussed in this chapter, it remains the most accurate, particularly when enumeration districts are small. In addition, there is increasing availability of high resolution digital orthophoto and high altitude CIR coverage for many urban areas in the U.S. that can be analyzed with the method. These data can generally be downloaded online from state or university GIS repositories without cost. In our experience, these data range around 30 to 15 cm in pixel resolution, are geocorrected to a map base, and are excellent quality. These characteristics make the interpretation of most dwelling diagnostic features a straightforward task.

These high resolution digital image data have some limitations. One limitation is the inability to view the photographs in stereo. Because of this, other clues such as shadows must be used to measure building heights. In addition, the end user has no control over the date of the photography. This not only includes year, but season as well. Figure 1 is a medium-density residential block near downtown Salt Lake City, Utah. The photography (originally in color) was acquired in late summer 2003 as part of a USGS program to acquire high resolution imagery of the most populated urban areas in the United States. The data are available to the public from the U.S. Geological Survey, EROS Data Center, Sioux Falls, SD (http://www.usgs.gov/).

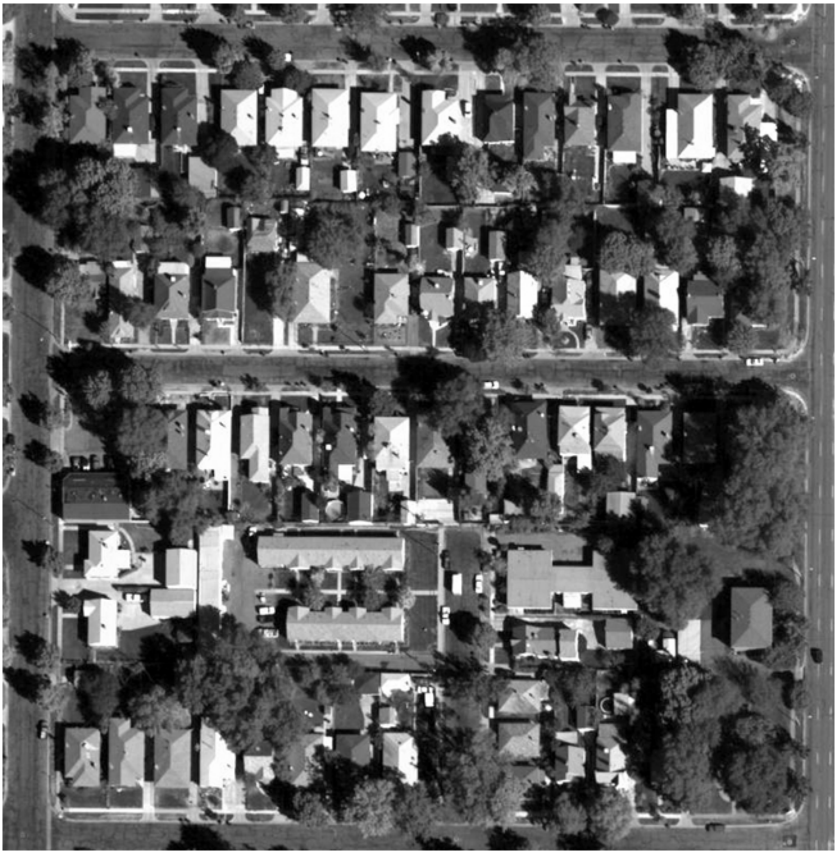

Figure 1. Leaf-on 30cm aerial photo of Salt Lake City, UT.

The pixel resolution is 30 cm. Figure 2 is a similar block in Pittsburgh, Pennsylvania, taken in leaf-off conditions. Classifying the buildings in the Pittsburgh imagery would prove a simpler task because the deciduous trees are not obscuring dwelling yards, roofs, etc. In contrast, the interpretation of the leaf-on Salt Lake City imagery would be more challenging. Of course if tree cover were an important diagnostic feature in a landtype schema, leaf-on imagery may be preferred.

Adeniyi (1983) summarizes our viewpoint on the dwelling unit identification approach when writing "The results have revealed, in general, that remotely sensed data have the capability to provide timely, verifiable, and relatively accurate intercensal population data, based on uniform criteria at local, metropolitan, and regional levels" (p. 546). Adeniyi is equally perceptive regarding the method's greatest limitation when stating, "[Each]

Figure 2. Leaf-off 30cm aerial photo of suburban Pittsburgh, PA.

be valid only for the particular area under consideration. Consequently, there seems to be a need to formulate for each cultural area a suitable model based on relevant attributes of the area" (p. 546).

4.2.2 Landtype Surrogates

Along with the lack of generality in interpretation keys, another drawback to counting dwelling units is the intensive labor required to do the photo interpretation. The method also requires the use of large scale photography and is not suitable for use with satellite imagery having a pixel resolution coarser than about 1 meter. Because of these shortcomings, the use of landtype zone as a population surrogate was developed. The premise of the approach is that landtype is related to housing densities which are in-stage coupled to population density.

Two variants are common. In the first variant, a schema of landtypes is developed and a characteristic population density for each type is assumed, measured, or estimated. The landtype zones are delineated on overhead imagery. Given that a district can contain different landtypes, the population for a district is next estimated in three steps. First, the area of each landtype in the district is determined. Landtype area is then multiplied by the population density assumed for the category. The resulting product is the total population for that landtype within the district. Summing those totals across all the landtypes in the district produces the total district population. In the second variant, the imagery of the enumeration districts is investigated to determine the different landtype percentages constituting the district. Those percentages become carrier variables in transfer equations relating the landtype constituent amounts to the district population density.

The need for accurate population data with which to calibrate the landtype → population transfer functions is the logical equivalent of calculating the average number of people per dwelling unit type in the structure counting approach. It is generally required for both variants.

Kraus et al. (1974) were some of the first researchers to advance the landtype surrogate approach for population estimation. In their experiment, four cities in California (Fresno, Bakersfield, Santa Barbara, and Salinas) were chosen for study. The goal of the research was to estimate the cities' population from high-altitude aerial photography with scales of 1:600,000 (panchromatic), 1:120,000 (CIR), and 1:60,000 (CIR). The interpretation schema utilized only four land use types; single family residential, multi-family residential, trailer park residential, and commercial / industrial. The entire built-up area of the four cities was placed into one of those four classes based on landscape appearance in the photographs. The area of each land use type was measured with a polar planimeter. "In order to obtain characteristic spatial population densities for the three residential land use categories, 1970 U.S. Census Block Data was used. Areas of a single residential land use were identified from the land use maps and located on Census Block data maps. Random samples of blocks within each residential land use category were then obtained to determine population densities per square kilometer for that land use. The spatial population density figures derived from each random sampling were then averaged to obtain characteristic spatial population densities for each residential land use category within each city" (p. 39). The total population for one of the mapped zones was the product of the characteristic population density and zone area. Summation of the zonal population across the city produced the total population estimate.

The results of the zonal procedure were a 7.0% population overestimation in Santa Barbara, and an average 7.2% underestimation in the other three cities. Two causes of the underestimation in Fresno, Salinas, and Bakersfield were given. The first was the inability to identify residences in older built-up business districts. Secondly, the enlargement of the original photography for easier interpretation unexpectedly increased the difficulty of identifying isolated individual apartments, causing an underestimation of area in the multifamily residential class. The overestimation in Santa Barbara was primarily due to the large lot sizes in many single-family residential zones – the characteristic population density applied to those zones in Santa Barbara was too large. To ameliorate these problems, Kraus et al. recommended; 1) a correction factor for "hidden" residential uses in commercial districts, 2) the use of larger scale air photography, and 3) a refined residential land classification system that permitted fine tuning of characteristic zonal population density factors.

In research reported by Adeniyi (1983), the objective was to examine the feasibility of systematically estimating Nigerian population with aerial photography. The research was warranted by the historic failure to accurately estimate Nigerian population using traditional methods, a failure attributed to the paucity of accurate social and administrative data. Additionally, urban planning demands of Nigeria required population estimation for small areas (e.g., voting districts) undergoing rapid urbanization. Based on significant preceding research (Adeniyi 1976, 1980) the project began with the simple hypothesis that population estimation based on land use zones would be appropriate for Nigeria. Two reasons were given for the use of the landtype zone method in preference to the dwelling unit method described above, both related to the communal housing structures used in Nigeria. First, the individual dwelling units were not readily countable on the available air photographs. Second, because each single storied structure might house between three and ten families of considerable and variable size, the utility of an "average family size factor" was dubious and would have produced significant estimation error.

Adeniyi chose the Federal Capital of Nigeria (Lagos) as the study site. The methodology was complex, required several steps, and will only be summarized below. Initially, different residential areas were delineated and classified on the 1:20,000 panchromatic photographs acquired in 1974. Housing quality and other sociocultural information was next collected from the air photographs. These photographically measured variables included building density, plot size, layout, garden existence, number of stories, dwelling type (e.g., apartment, communal), and building usage.

Examination of these variables suggested a landtype schema with nine species. Once the schema was completed, the analysis required that the population densities of each residential landuse category be gauged. Using a random sampling scheme stratified by landuse category, information on family size and number of families for each residential structure type was determined by limited field survey. The total number of field samples (i.e., residential blocks) was 58 with the number of samples per land use ranging from 1 to 20. A total of 3,479 buildings were included in the 58 blocks. A cluster analysis of these 58 field samples using the field data alone regrouped the field samples into nine temporary subsets for the purpose of exploring intraresidential and interresidential class differences and similarities. The clustering exposed two broad divisions in the land zones. The first division was the planned land zones – planned residential areas with apartment housing of moderate density. The second division consisted of the higher density communal dwelling structures considered traditional and unplanned. It was also observed that many of the residential landtypes could be distinguished almost completely by the density of residential buildings typifying them.

Average population density figures were calculated from the field data for each residential land use. Regression analyses were then used to optimally model the population density of each land use using all of the survey and air photograph variables as candidate independents. For the complete sample of 58 blocks, three variables were able to explain 90% of the variation in the population density; 1) density of communal type buildings, 2) average population per building, and the 3) density of all buildings. Different residential land use classes had different models; however the two variables appearing the most frequently in the models were density of communal type buildings and average population per building.

Because the data gathered had been exhausted in the model building phase, validation of the models was impossible, but Adeniyi made the following observations.

1. The Lagos population estimates in zones of planned residential development were quite accurate, primarily because the average population density value for the planned residential land uses was uniformly applicable to all samples of that type. If air photos were used operationally to estimate Nigerian population, fieldwork to establish the average building population for planned zones would be minimal.

2. In contrast, fieldwork to support estimates for the unplanned residential landuse categories would be necessarily extensive. The highly variable average population density factor would require tailoring for different regional areas.

Another example of the residential landtype approach with a simpler methodology is Olorunfemi (1984). This study was conducted in the city of Ilorin, Nigeria, a city of ½ million population that serves as the capital of Kwara State. The goal of the study was to "define a mathematical model which may be used in conjunction with data on housing land area measured from aerial photography to obtain urban population estimates for Nigerian cities" (p. 221). The photographs for the study were acquired in 1950 and 1963 at scales of 1:2400 and 1:12000 respectively. Census data to support the research was taken by survey method in 1952 and 1963.

A total of 74 square sample sites of 4 hectares each were randomly selected from topographic maps and their location transferred to the two sets of aerial photography. The area (percentage) of each major landtype within each sample site was measured using a dot grid. The landtype classification schema used six categories; 1) indigenous residential type housing, 2) barrack / flat housing, 3) flat housing, 4) uncompleted housing, 5) bare ground / grass / agricultural land, and 6) trees. The population of each sample site was determined from the census data. Multiple regression analysis was conducted to model population within the 4 hectare sample sites as a function of the area in each of the six categories.

For the 1950 case, population was significantly correlated with the percentage area of uncompleted housing (*r* = 0.66)[5], the percentage of trees (*r* = -0.66) and percentage area devoted to barrack / flat housing (*r* = 0.57). A linear regression model using all six variables explained 72.6% of the variation in population density. For the 1963 data set, the only significant explanatory variable was the percentage of area devoted to flat housing (*r* = - 0.51).

Olorunfemi concluded there was a functional relationship between landtype and population density which justified the use of landtype as a population surrogate. The goodness of the regression models was deemed sufficient to warrant an examination of its utility for wider application in Nigeria. "It should be stressed, however, that, for this method to be useful in generating nationwide data, there is need for further research aimed at testing the applicability of the model in cities with similar and/or different characteristics" (p. 227). Olorunfemi also considers the method particularly appropriate in communities where housing land area is known beforehand or population data is unavailable because of "remoteness, political obfuscation, or insufficient resources to conduct frequent census enumerations" (p.227).

Since the launch of Landsat-1 in 1972, the mapping of landcover from medium resolution satellites has become operational in many disciplines such as range management and agriculture. The potential for adapting the landtype zone population estimation method from aerial photography to satellite imagery was natural. The work of Langford et al. (1991) serves as an excellent example of this adaptation. One objective of this study was to model the 1981 population of 49 wards in four districts of northern Leicestershire, England using the landtype surrogate approach. The methodology was straightforward. First, using automated image processing methods, a satellite image dated July 1984 was classified into various landtype categories on a pixel-by-pixel basis. The single satellite image covered all 49 wards of the study area. Creation of the landtype map proved challenging, primarily because the unsupervised classification highlighted land cover differences in the rural hinterlands of the area but did not sufficiently discriminate between important urban landtypes. After some trial and error, principal components analysis of the original seven band TM image resulted in three principal component bands that revealed urban land cover differences necessary for accurate discrimination. The resulting map consisted of 12 landtypes collapsed into five broader categories; 1) commer-

[5] Unless otherwise noted, correlation coefficients (*r*) refer to Pearson's r.

cial and industrial, 2) high-density residential, 3) ordinary residential, 4) quarries, woods, water bodies, and 5) agriculture. Once the landtype map was complete, the 1981 census ward boundaries were digitized, rasterized, and imposed on the digital landtype map. This permitted the area of each type (listed above) to be tallied for each of the 49 wards. Except for agriculture, the greatest difference in cover amounts between the various wards was in the ordinary residential category – ordinary residential land cover within a ward varied from 10 to 430 hectares with a mean and standard deviation of 158 and 95 hectares respectively.

Simple correlation analysis revealed that total ward population was most highly correlated with the ordinary residential (r = 0.75) and commercial / industrial (r = 0.60) land area in the ward. Total population was more weakly correlated with agriculture (r = -0.28) and high-density residential (r = 0.33) area. Encouraged by the results of the correlation analysis, several regression models were created that explained total ward population as a function of the five independent variables listed above. The models differed primarily on the number of variables used, the use (or not) of a Poisson error term, the fitting (or not) of an intercept, and whether negative coefficients were permitted. The last requirement (no negative coefficients) was designed to preclude models that might generate negative population counts. Regression equations that had a non-zero intercept were also considered logically flawed.[6] In summary, it was "argued that that any statistical model linking pixel counts of land cover to population should be simple, linear, additive, and without any intercept constant" (p. 67). The most effective model produced by Langford et al. included only two variables, ordinary residential landcover and high density landcover area. This ordinary least squares model with only additive coefficients produced an R^2 of 0.82. Areas of underestimation and overestimation, sometimes severe, were noted by spatially analyzing the residuals from the regression. The residual patterns from the different models were similar.

Comments

Although the required methodology differs slightly, the use of landtypes as population surrogates is equally applicable to air photos and satellite data alike. The use of satellite data has some particular advantages. Unlike analog air photography, satellite imagery lends itself to automated interpretation, classification, and georeferencing. For large areas with no pre-

[6] Harvey (2002a; p. 2086) discusses the issues of negative coefficients and zero intercepts in some detail and provides alternative opinions.

existing air photo coverage, purchasing satellite data is much less expensive than contracting for an aerial photography acquisition mission. For example, one Landsat image covering about 34,000 km^2 can cost less than $1000 U.S. Because of its lower expense, satellite imagery can be acquired more often (or on demand) and can thus facilitate more timely population estimates than air photography. Although high resolution satellite data are available, the landtype surrogate approach only requires less expensive moderate resolution imagery (i.e. 30m - 50m pixel resolution). Satellite data can also provide a fine degree of spatial granularity over large urban areas.

The landtype surrogate approach does have disadvantages. It is less accurate than counting dwelling units. Although a landtype classification schema may only contain five or six categories, the task of creating the landtype schema is critical for success. As illustrated by Langford et al. (1991), the task of generating an urban landtype map with sufficient detail to support population estimation may require substantial trial and error. Nonetheless, the use of landtype surrogates continues to be an important tool in geographic urban analysis.

4.2.3 Pixel-based Estimation

Pixel-based estimation is an approach designed entirely for moderate resolution satellite imagery. In its basic form, the goal of pixel-based estimation is to model population or population density directly as either a function of multiband satellite sensor reflectance values or some mathematical derivative thereof.[7] Adopting the logic of Iisaka and Hegedus (1982), the justification for this approach lies in the nature of the pixel itself. In an urban area, a single satellite pixel will contain a variety of land cover types that contribute to the spectral reflectance of the pixel. Figure 3 shows an area of several pixels covered by a Landsat MSS image compared to an aerial photo with 6 inch resolution to demonstrate the variety of objects within a single 79 x 79 meter pixel. Exemplar cover types are rooftop shingle, road surface concrete, lawn grass, and parking lot asphalt. As population density varies, the relative percentage of these cover types covaries. This variation accordingly modulates the spectral signature of the pixels having an urban footprint (Iisaka and Hegedus, 1982). While the causal relationship may be <u>population</u> → <u>housing</u> → <u>landtype</u> → <u>spectral</u>

[7] The term "spectral features" is given to these mathematical derivations. It should not be confused with "spatial features." Example spectral features include vegetation indices, principle components, and texture measures.

signature, the pixel based estimation usually proceeds spectral signature → population with the landtype and housing treated implicitly.

Figure 3. Landsat MSS image compared to aerial photo of Provo, Utah.

The goal of Iisaka and Hegedus (1982) was to model the relationship between spectral reflectance and the population of metropolitan Tokyo. The satellite images acquired from Landsat 1 and Landsat 3 were dated November 1972 and January 1979 respectively. Census data required to calibrate and validate the modeling were acquired in 1970 and 1975. One of the assumptions required in modeling population as a function of spectral signature "is that the environmental alteration…should share similar characteristics in different areas, in both quantitative and qualitative respects" (p.261). This permits measurements made in sample areas to be logically applied to other locales. The authors claimed that the homogeneity of housing materials, dwelling size, land use systems, and housing density in residential Tokyo satisfied this assumption.

A total of 88 sample sites (25 hectares each) outside the Tokyo central business district were initially identified. The size of the sample sites (500 m × 500 m) corresponded to the resolution of the government census maps of the Tokyo ward area. Once identified, the population for each sample site was extracted from the 1970 and 1975 census maps. The satellite imagery was resampled and georegistered to the census maps and the mean spectral values for the four multispectral scanner (MSS) bands determined for each of the 88 sample sites. This created a data set of 88 records having five fields each. Population was the dependent variable whereas the carrier variables were the four MSS bands.

Exploration of the data indicated that the green and infrared bands were strongly correlated (linearly) with population. Regression analyses were conducted to determine whether population for the sample sites could be explained as a function of the four-band spectral signatures. As expected, regression equations utilizing the green and infrared bands were most capable of predicting population density. The signs for the coefficients remained the same for the 1972 and 1979 data although the magnitudes of the coefficients were different. Examination of the regression equations showed reflectance in the green band increasing and reflection in the two infrared bands decreasing with increasing population density. Iisaka and Hegedus do not explain the physical basis for the signs of the coefficients or the reason why they are different between the two years. Based on our own research in North America, we cautiously suggest that denser urban build-up associated with greater population densities co-occur with less urban vegetation, hence the inverse relationship with infrared reflectance.[8] At first blush the same argument might also suggest the same inverse relationship between population and green reflectance. However the increased green reflectance from concrete and other lightly colored inert materials more than compensates for the loss of green reflectance from sparse urban vegetation in such situations. Likewise, we suggest that the difference in the reported regression coefficient magnitude between the two years *might* have been reduced by employing radiometric / atmospheric correction and standardization methods which have become common in Landsat data processing since that time (e.g. Singh, 1989).

Overall, the models generated multiple R values of 0.84 and 0.77 for the 1972 and 1979 studies respectively. Judicious removal of a few atypical sites improved the multiple R values to 0.94 (n = 60) for 1972 and 0.90 (n = 62) for 1979. Nonconformant sites were residential sample sites containing train stations, schools, churches and other features not prevalent in the Tokyo residential area.

The next important milestone of intraurban population estimation research was reached by Lo (1995) who rigorously evaluated the use of SPOT im-

[8] An increase in shadow from increased urban "canyonization" is an alternative explanation.

agery for population estimation[9] using methods predicated on Iisaka and Hegedus' (1982) earlier success. Whereas the Landsat MSS used by Iisaka and Hegedus (1982) utilized four spectral bands with a pixel size of 79 m, the SPOT imagery used by Lo had three bands (i.e., green, red, infrared) and a pixel resolution of 20 m. When the project was initiated, it was thought that the smaller SPOT pixel size would be a decided advantage, "Because of the low spatial resolution of the Landsat-MSS data, the spectral radiance is the average of reflectance of different cover types over an area of 79 m by 79 m in a pixel. The spectral reflectance of the residential cover, on which population estimation has to be based, is therefore highly diluted. This dilution will likely affect the accuracy of the population estimates. … Because of its better resolution, each SPOT image pixel covers a much smaller area on the ground, and hence the spectra radiance is more representative of its ground cover characteristics than the 79-m Landsat-MSS counterpart" (p. 18). In the exploratory work of Iisaka and Hegedus (1982) cited above no attempt was made by the authors to incorporate any *a priori* knowledge about the study area that might permit different regression models to be used in different neighborhoods. In contrast, the objective of Lo was to model population density in a metropolitan Hong Kong study area (i.e., Kowloon) as a function of SPOT spectral reflectance while employing GIS technology to permit different regression transfer equations to be used with different landtypes.

As mentioned by Lo, the mixed landuse in Kowloon was a significant challenge. Landuse was complex with residential areas intricately mixed with non residential areas. Transitions between low-density residential areas and high-density overcrowded areas were abrupt. Multistoried buildings in Kowloon not only housed multiple dwelling units, but the buildings themselves were multiple-use, serving commercial and industrial functions too.

Population data for Kowloon used to support the research was collected by the Hong Kong Census and Statistics Department in 1986 for 60 planning units via complete enumeration. Because of computer storage limitations, only 44 of the planning units were examined in the study. These ranged in area from 6 to 291 hectares. The SPOT data used for the population estimation was acquired in January, 1987. After significant preprocessing, the

[9] Lo (1995) also treated the counting of dwelling units with SPOT imagery. To simplify the review, we cite only the results for the population estimation component of the research. The dwelling unit estimation results closely paralleled those of the population estimation.

SPOT data was georegistered to the boundaries of the planning units. This permitted mean spectral reflectance values to be calculated for each unit. By calculating the area of each planning unit with the GIS, population densities for each were also determined.

Exploratory analysis indicated a moderate negative correlation (r = -0.62) between mean infrared reflectance and population density. After these initial findings, several regression analyses were performed to model planning unit population as a function of mean planning unit spectral reflectance. Twelve of the planning units were manually selected to calibrate the regression equations and the remaining 32 were reserved for testing regression goodness. Regression models were built using all three bands as well as simplified regression models using only the infrared band. The dependent variable in the regression model was population density rather than population counts.[10] The model using three bands was capable of estimating population of the whole study area with a relative error of 1.7%, whereas the infrared band model resulted in an error of 15.0%. When population estimation was attempted at the smaller scale of the planning unit, serious estimation errors sometimes exceeding 500% were encountered, primarily in commercial and industrial planning units. The mean relative error for the micro-scale planning unit estimation was about 75%.

Because each planning unit in both Kowloon and the greater Hong Kong metropolitan area was a mixture of both residential and nonresidential land uses, Lo sought to refine the population estimation process. The refinement required that the pixels actually representing residential land use be identified within each planning unit. This would permit the regression equations to be applied to those residential pixels alone and avoided the errors associated with attempting prediction for those pixels known to be predominantly commercial or industrial.

This refinement was completed by classifying the SPOT image into eight landtypes which included both low-density residential and high-density residential categories. An average per-pixel population density for high and low density residential zones was calculated using the census data for the 12 calibration planning units. The refinement produced a modest decrease in the population estimation errors of the smaller planning units, and the absolute mean relative error dropped to 67%.

---

[10] When multiplied by the area of a single SPOT pixel (0.04 hectares), the population per pixel could be easily estimated from the density.

For its rigor in methodology and actual success in modeling population density of small districts, Harvey (2002a) represents a landmark work of significant dimensions that should be read in its entirety. Harvey's goal was to model small-area population densities for Australian census collection districts (CDs) using spectral features measured from Landsat TM imagery.

Imagery of Ballarat Statistical District (west of Melbourne Australia) containing 138 CDs was used to build models of urban population density. Thematic Mapper imagery of Geelong Statistical District (225 CDs), nearly 100 km southeast of Ballarat was then used for model validation. Population data supporting the study was collected in 1986 by the Australian Bureau of Statistics and was preprocessed to correspond more closely to the February 1988 imagery date.[11]

Recognizing that "one obvious problem was that the values of the dependent variable ranged over three orders of magnitude" (p. 2079), both logarithmic and square root transformations of CD population density were calculated before multiple regression modeling began. A set of 80 predictor variables was submitted to a host of stepwise regression analyses in an effort to find the best predictive variable subset. These included the following classes of transformations, calculated for each CD by using the pixels captured within the digitized boundaries of the CD:

- Mean TM band reflectance
- The square of the mean of TM band reflectance
- The cross product of the mean TM band reflectance
- The ratio of mean reflectance for two TM bands
- The difference-to-sum ratio for two TM bands
- The TM band variance
- The TM band standard deviation
- The TM band coefficient of variation
- Normalized bands

The following spectral transforms were calculated on a per-pixel basis and then summarized for each CD by calculating means and measures of variation. These were numbered among the 80 multiple regression variables.

- Selected normalized bands
- Selected band ratios
- Difference / sum ratios of selected bands

[11] See Harvey (1999) for a discussion of this preprocessing.

- Hue transforms of selected bands using both rectangular and cylindrical coordinates.[12]

Summarizing Harvey's voluminous results, model R^2 values of about 0.90 were obtained with model subsets containing between four and nine variables. The best predictors were the mean and standard deviation of per-pixel spectral features listed above. The dependent variable transform associated with the highest R^2 regressions was the square root rather than logarithmic transform.

For validation, six multiple regression models predicated on the Ballarat study area were applied to the TM imagery of the Geelong Statistical District. Population counts and densities for the 225 Geelong Statistical District CDs were thus estimated and then compared against the true CD population counts. The two models based on band means alone produced very poor total population estimates. The model based on the per pixel measures listed above produced the best total population estimate in the validation CDs and had an urban total underestimation of only 3%. The following observations were made:

1. Models of high complexity performed better than simpler models.

2. Populations of lower density in rural areas were consistently and seriously overestimated. The errors were not large in terms of population numbers, but rather in percentages of the true population counts. Regarding this rural overestimation, Harvey comments, "It is concluded that the potential of this methodology is limited by heterogeneity of both land cover and population density within the individual CDs, and that are, in principle, unlikely using this approach. In particular, the sacrifice of detailed spatial information leaves no way to respond to the problem of over-estimation of population in large areas of low density" (p. 2093).

3. Given that models driven by per-pixel spectral indicators were superior to those calculated for the CDs (e.g., CD band means), Harvey conjectures that "models formulated [at a lower] spatial level can produce relatively accurate…population estimates for larger spatial aggregates, but not for spatial units at the same level of aggregation" (p. 2093). Harvey concludes that future modeling

[12] See Jensen (2005; pp. 164-167) for a discussion of hue transforms.

> should be logically done at a lower level of aggregation than the CD to minimize errors in population prediction. He also points out the difficulty operationalizing this within his Australia study area – while spectral data may be available at the smaller pixel level, the census calibration data is available only at the larger level of the CD, and thus precludes lower level modeling. Methods to overcome this obstacle by exploiting an expectation-maximization statistical algorithm are presented in Harvey (2002b).

The primary objective of the research conducted by Li and Weng (2005) was to develop and compare methodologies for estimating the population density of Indianapolis, Indiana using Landsat ETM+ data. As a justification for their research, Li and Weng claimed that previous research "rarely [had] explored the integration of spectral, textural, temperature data, and advanced transformed remote sensing variables to estimate population. [13] Such incorporation may provide new insights for population density estimation" (p. 948). The ETM+ satellite data were acquired on 22 June 2000. The population data, based on census blocks, were obtained from a GIS vendor and aggregated into census block groups (CBGs).

Li and Weng's research objective required that several spectral features be examined to determine their correlation strength with population density. These spectral features included; 1) the first principle components of the ETM+ visible and optical infrared bands, 2) six different vegetation indices, 3) variance images (with various local window sizes) calculated from ETM+ red and middle-infrared bands, 4) surface temperature from ETM+ Band 6[14] and 5) impervious surface and green vegetation fraction images produced from decomposition of the six ETM+ visible and optical infrared bands.[15]

The study area consisted of 658 CBGs. An initial investigation required 162 samples for model building and the remainder for model validation. Like their predecessors, the samples used for creating the models were not entirely chosen at random. In this Indianapolis study, all the block groups with low and high population densities were included among the 162, whereas the medium population density CBGs were sampled randomly.

[13] Harvey (2002a) is an obvious exception to this generalization.

[14] See Weng et al. (2004) for detailed information on how surface temperature was derived.

[15] See Lu and Weng (2004) for more explanation about the fraction images.

Simple correlation analysis was used to explore the relationship between population density and the spectral features mentioned above. The simple Pearson's r showed only a weak correlation between population density and the spectral bands, principle components, vegetation indices, fraction data, and texture – r never exceeded 0.4 in absolute value. The correlation between temperature and population density was moderate (r = 0.519). It is notable that while the correlation between population density and the middle-infrared band was nearly zero, the correlation between the texture calculated in that band with a 7 x 7 window was substantive (r = -0.402). It was also found that transformations of the dependent variable (i.e., natural logarithm, square root) improved correlations moderately. The best multiple regression equation using a subset of the spectral features generated an R^2 of 0.83. A residual error map showed the greatest misestimation occurring in CBGs of extremely high and low population density.

According to Li and Weng, "In order to improve population estimation results, separating the population density into sub-categories such as low, medium and high densities, and developing models for each category" (p. 952) was deemed a necessity. However, the results of using separate regression models for the three strata of population density were mixed. In general, stratification improved the results, nonetheless, with R^2 values for low and high density population density CBGs never even reaching 0.2, Li and Weng questioned whether Landsat ETM+ data was suitable for modeling extremes in population. For medium density population, the results were much better with R^2 approaching 0.90. The best predictor variables for the medium density models included red band texture (7 x 7 window), thermal temperature, the simple ratio of the near infrared and red bands, the transformed normalized difference vegetation index,[16] the soil adjusted vegetation index,[17] infrared reflectance, and the value of the first principle component image. Although misestimations for low and high density population CBGs were still significant the total population estimate error for the Indianapolis study area was only 3.2%.

Comments

We consider pixel-based estimation insufficiently explored to provide a generalized judgment of its worth. Because of differences in urban physiognomy, the results of the urban Tokyo, Ballarat, Hong Kong, and Indianapolis studies are difficult to apply to other worldwide cites. However, we

[16] See Deering et al. (1975)

[17] See Huete (1988)

agree with Li and Weng (2005) that "using remote sensing techniques to estimate population density is still a challenging task both in terms of theory and methodology, due to remotely sensed data, the complexity of urban landscapes, and the complexity of population distribution" (p. 955).

The foregoing review suggests several things. First, practitioners using per-pixel estimation should expect estimation problems in areas of extremely low or high population density. Second, per-pixel estimation using spectral radiance can benefit from stratification of the landscape that permits different regression equations to be built for specific landtypes. Third, because population errors of underestimation and overestimation tend to be compensating, accuracy of the approach will increase proportionally with the area of the enumeration units. Fourth, efforts to define spectral features more related to population density than reflected spectral radiance are warranted. The use of texture by Li and Weng (2005) and the spectral measures of Harvey (2002a) have already been mentioned. Webster (1996) presented several other texture measures appropriate to urban population modeling, all of which can be automatically derived from satellite imagery. In addition, indices derived from spectral reflectance commonly used in vegetation analysis[18] may likewise be more closely related to population density than spectral reflectance values used alone. Other indices specifically related to population density may require development. Finally, as illustrated by Li and Weng (2005), urban temperature measured by satellites is modulated by the amount of inert or built-up land within the thermal sensor footprint. Further exploring the use of thermal temperature as a surrogate for population density might likewise prove fruitful. If so, then the use of imagery from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) with its five thermal bands may provide improved estimates of population density than those possible with ETM+.

4.3 Case Study

The goal of this preliminary case study was to determine whether the pixel-based approach to population estimation developed by Iisaka and Hegedus (1982) and extended by Lo (1995) would be successful in modeling population density in the Wasatch Front of Utah. Three specific objectives were outlined.

[18] See Jensen (2007; pp. 382-393) for a review of popular vegetation indices.

1. To determine whether the population density among 1990 census block groups (CBGs) could be modeled as a function of the spectral characteristics of the census blocks as measured by Landsat Thematic Mapper data acquired in 1990.

2. To determine whether other spectral measures such as texture and the Normalized Difference Vegetation Index could model population density more accurately than simple spectral reflectance used alone.

3. To determine whether thermal data collected by Landsat Thematic Mapper could be effectively used to model population density in the study area.

This case study is a small part of a larger project to determine whether the pixel-based approach can be used to create accurate intercensal population estimates of the same Wasatch Front region.

Study Area

Demographics. As shown in Figure 4, the study region includes the major metropolitan areas of Utah. Over a dozen incorporated entities are part of the study area, including Bountiful, Salt Lake City, Taylorsville, Sandy City, Orem, Provo, American Fork, and Spanish Fork. It is an area of steady population increase. The average population growth along the Wasatch Front between 1970 and 2005 was about 2.5% per year.[19] The population growth has been steady primarily because of Utah's historically large family size and relatively high fertility levels. These factors make Utah somewhat unique in the U.S.; the majority of its population growth (80%) is from natural increase (GOPB 2005). Even when net migration is low or negative, Utah still experiences population growth driven by natural increase. Nonetheless, migration will continue to be an important factor in Utah population growth. Over the last fifteen years the Wasatch Front has averaged about 25,000 new residents a year and the GOPB predicts that about 42,000 new residents a year will make the Wasatch Front region their home between 2005 and 2030. Given the rapid and steady population growth, as well as its concomitant affect on the housing and construction sector and land use change, the Wasatch Front region is an excellent case study site.

[19] This is slightly below the 2.6% average annual population growth experienced by other states in the Mountain West census region, but significantly higher than the U.S. annual average population growth of 1.1%

Figure 4. Municipalities in the study area along Utah's Wasatch Front.

Landtypes. Based on research conducted primarily in Salt Lake City, Utah, Ridd (1995) categorized urban fabric into three components; 1) vegetation, 2) impervious surfaces, and 3) bare soil. The common abbreviation for this triad is VIS. Impervious surfaces include concrete, asphalt streets, asphalt roofing, shingles, and metal roofing. Vegetation includes grass, tree, and shrub categories (after Ridd, 1995).

Using Ridd's model, Hung (2002) extensively studied the VIS components in Salt Lake City, findings we consider generally applicable to the whole study area. Hung found that urban commercial districts in Salt Lake City are strongly dominated by impervious surface (86%). Low density residential (i.e., single-family residential) zones are a mix of vegetation and impervious surface with vegetation representing nearly 70% of the land cover. Change from low through moderate to high density residential zones showed an increase in impervious surface area (24% → 34% → 46%) at the expense of vegetation (69% → 56% → 43%). Soil was a large component only in industrial zones (24%) and was about equal to vegetation coverage (26%).

Data and Methods

Two data sources were required to conduct this study. The first source was the 1990 Decennial Census of the United States. The specific variable extracted from the census was total population aggregated by census block groups (CBGs). A total of 812 CBGs were initially part of the study area. The average CBG size was 3053 hectares with a range from 11.3 to 520,000 hectares. Figure 5 is a map of population density in 1990 for the study area. The average population density among CBGs was 16.8 people / hectare (s = 12.6) and the average density of housing units was 6.3 units / hectare (s = 5.8). The densest population and housing recorded among the CBGs was 110.4 people / hectare and 48.5 units / hectare respectively. These high densities are found in areas of student housing adjacent to Brigham Young University in Provo, Utah.

The second data source was a Landsat ETM+ image dated May 28, 2000 (Figure 6). Table 1 shows the fundamental characteristics of the Landsat TM sensor. The quality of the image was good, but cloud cover masking was required to avoid problems relating spectral signature to population density -- portions of CBGs containing clouds were removed from the analysis. Unfortunately, thin smoke also partially obscured a few of the CBGs. Since this smoke was found at the interface between the suburban and rural areas, the affected image areas were not eliminated from the study. Instead, the areas were manually delimited and the contrast and brightness adjusted until they matched the surrounding area. This manual approach was only partially successful. The image data was radiometrically corrected, and standardized to reflectance/emittance as measured at the sensor.

Recognizing from the reviewed literature that derived spectral features may be more predictive of population density than simple spectral reflectance alone, several derived variables were generated from the seven TM bands for each CBG.

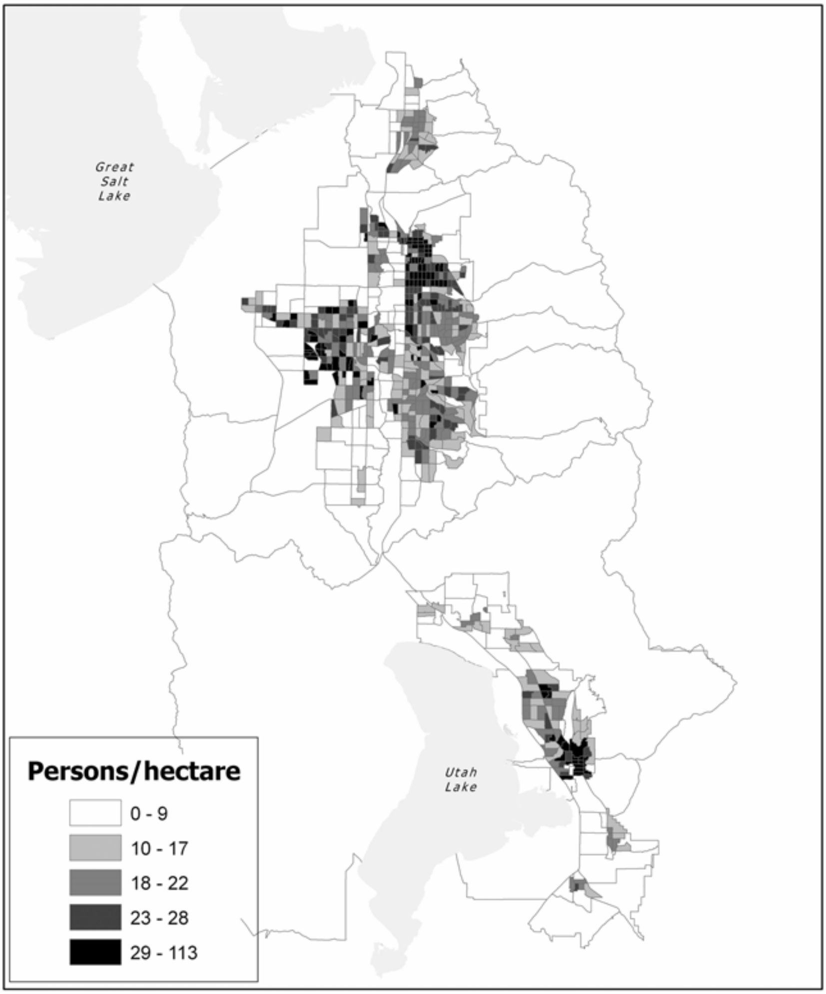

Figure 5. Population density of census block groups in the study area.

Figure 6. Band 4 (IR-1) of a Landsat ETM+ image acquired of the study area May 28, 2000 (2σ contrast stretch).

Table 1. Fundamental characteristics of the Landsat TM sensor. Columns 1, 3, and 4 from Landsat 4 Data Users Handbook (1984).

Band Number	Band Name	Band Wavelength (micrometers)	Nominal Resolution (meters)
1	Blue	0.45-0.52	30
2	Green	0.52-0.60	30
3	Red	0.63-0.69	30
4	IR-1	0.76-0.90	30
5	IR-2	1.55-1.75	30
6	Thermal	10.40-12.50	120
7	Mid-IR	2.08-2.35	30

The final variable set used in the correlation and regression analyses included the following:

- Mean Blue reflectance (B_{μ})
- Mean Green reflectance (G_{μ})
- Mean Red reflectance (R_{μ})
- Mean IR-1 reflectance ($I1_{\mu}$)
- Mean IR-2 reflectance ($I2_{\mu}$)
- Mean thermal brightness temperature (T_{μ})
- Mean Mid-IR reflectance ($I3_{\mu}$)
- Standard deviation of Blue reflectance (B_{σ})
- Standard deviation of Green reflectance (G_{σ})
- Standard deviation of Red reflectance (R_{σ})
- Standard deviation of IR-1 reflectance ($I1_{\sigma}$)
- Standard deviation of IR-2 reflectance ($I2_{\sigma}$)
- Standard deviation of thermal brightness temperature (T_{σ})
- Minimum of thermal brightness temperature (T_n)
- Maximum of thermal brightness temperature (T_x)
- Range of thermal brightness temperature (T_r)
- Standard deviation of Mid-IR reflectance ($I3_{\sigma}$)

The standard deviation, range variables, maximums and minimums were designed to be rough measure of spectral texture[20] in the CBG.

After exploratory correlation analysis, stepwise multiple regression analysis was utilized to build regression models explaining population density in

[20] Several popular texture measures, including fractal measures, were tried, but all proved less effective than these simple statistics.

the CBGs as a function of the variables listed above. Following the pattern of Harvey (2000a), several issues were considered in judging the goodness of a regression. The magnitude of the adjusted R^2 was considered. If all else were equal, a regression equation producing a higher R^2 was preferred. Simplicity was a second consideration. When regression equations producing similar R^2 values were compared, the equation using the fewest predictor variables was deemed best. Logical consistency was also necessary. Equations likely to produce negative population densities were discarded. Equations with variable combinations that appeared illogical were discarded. For example, it made little sense to include both $I1_{\mu}$ and $I2_{\mu}$ in the same equation unless we could reasonably explain why the addition of the second band added explanatory power not contained in the first. Equations exhibiting multicollinearity symptoms in the stepwise regression process were also discarded. Although a logarithmic transform of population density was required to improve linearity, we otherwise avoided variable transformations and higher order terms.

Results

Table 2 shows the simple correlation between the natural logarithm of population density (P_{ln}) and the independent variables listed above. The correlation between population density and spectral reflectance follows the same pattern reported by Li and Weng (2005, Table 3). Correlations are low but explainable patterns emerge. Blue, Green, Red, and IR-1 reflectance increase with increasing population density. Thermal temperature also increased with increasing density. In contrast, Mid-IR reflectance and IR-2 reflectance are inversely related to population density.

In the context of Hung (2002) summarized above, we consider these relationships largely a function of the relative amounts of vegetation and impervious material within the CBG. As housing unit density increases, the amount of concrete, asphalt, shingle, and other nonporous surfaces increases with the concurrent loss of grass, trees, and shrubbery. Since impervious surfaces reflect a larger proportion of incident visible light than vegetation does, the increase in reflectance with increased housing density is logical. Our interpretation of the similar increase in IR-1 reflectance is less obvious. Vegetation has a high IR-1 reflectance. Our tentative explanation for the sign in Table 2 is that the IR-1 reflectance increase from impervious surfaces more than offsets the loss of IR-1 reflectance due to vegetation loss.

Table 2. Pearson's *r* correlation between CBG spectral features and P_{ln}. Given the sample size of 807 CBGs, all are significant at the 0.01 level.

Spectral feature		*r*
Blue reflectance	Mean	0.202
	S.D.	-0.184
Green reflectance	Mean	0.185
	S.D.	-0.190
Red reflectance	Mean	0.117
	S.D.	-0.219
IR-1 reflectance	Mean	0.210
	S.D.	-0.203
IR-2 reflectance	Mean	-0.176
	S.D.	-0.427
Thermal temperature	Mean	0.238
	S.D.	-0.671
	Minimum	0.653
	Maximum	-0.488
	Range	-0.773
Mid-IR reflectance	Mean	-0.127
	S.D.	-0.352

Table 2 also demonstrates that spectral reflectance variability within a CBG is frequently a better predictor of population density than mean spectral reflectance. This agrees with the results of Harvey (2002) among others. This is clearly the case for the IR-2, Mid-IR and Thermal bands. In all cases, population density was inversely related to the variability features. We tentatively suggest that the increase in impervious surface associated with increased housing density is betrayed as an increase in urban surface homogeneity as vegetation amounts become more limited.

Given the weak correlations generally throughout the table, the strength of the thermal band features as predictors of population density is striking. This agrees with findings by Li and Weng (2005) for the Indianapolis, Indiana metropolitan study site. The reason for thermal temperature and population density covariance is well reported. Thermal temperature increases with the increased proportion of inert material associated with increased dwelling structure density. The high negative correlation between population density and temperature indicates that CBGs with higher population densities also have less temperature variability because the impervious material is ubiquitous throughout the CBG. Areas of lower population density will have built-up zones of housing (warmer) interspersed with parks, grass, empty vegetated lots, pastures, and agricultural areas (cooler).

Given that the relative amounts of vegetation and impervious material within a CBG were fundamentally responsible for the spectral reflectance → population density relationship, we generated some ratios to better represent the inverse relationship between vegetation and impervious material within the study area. These included:

- Mean CBG red band reflectance / Mean CBG IR-2 reflectance ($R_{3/5}$)
- Mean CBG red band reflectance / Mean CBG Mid-IR reflectance ($R_{3/7}$)

Multiple Regression Models. Table 3 contains the best regression models as judged by the criteria discussed previously. In all the models, the natural logarithm of the CBG population density (P_{ln}) was the dependent variable. All of the regression equations, constants and variables are significant at a 0.05 level minimum.

From the perspective of parsimony, it appears that the three-variable model would be preferred for practical application. This model produced a multiple R of 0.80 and an adjusted R^2 of 0.64. Models with more predictor variables produced models with R^2 values exceeding 0.70, but were difficult to interpret. As the three variable model shows, the best predictive combination includes temperature range, the first TM infrared band, and the ratio of the red band to IR-2 reflectance.

Table 3. Best regression equations to model P_{ln}.

Number of predictors	Equation	R^2 (adjusted)
1	$4.390 - 0.176\ T_r$	0.60
2	$2.389 - 0.173\ T_r + 6.65\ I1_{\mu}$	0.62
3*	$-0.164\ T_r + 8.685\ I1_{\mu} + 3.107\ R_{3/5}$	0.64
4*	$-0.161\ T_r + 9.362\ I1_{\mu} + 3.528\ R_{3/5} - 3.003\ I3_{\mu}$	0.64

* = the constant of the regression equation was not significant and is not included in the equation shown.

Residual Analysis. Figure 7 is a map of standardized residuals from the regression. Generally speaking, the best prediction was obtained in CBGs with moderate to small area in the central corridor of the study area. The relationship between CBG area and total population error[21] is linear and strong ($r = +0.853$, $n = 812$). These results are different from Lo (1996) but correspond to the results of Harvey (2002a). In the Wasatch Front, the large CBGs have boundaries which cross mountain slopes perpendicularly and capture areas of rural desert, agriculture, woodland and forest that form the hinterland of the core city area. Obviously the mean spectral values of these CBGs do not fairly represent the spectral character of the urbanized proportion of the CBG.

Any error in the population density produced by the regression equation also resulted in enormous estimation errors when multiplied by the area of large CBGs. For example, the largest CBG in the study exceeded 520,000 hectares, and had a true population of only 192. With very sparse vegetation, it has an average spectral signature similar to high density residential zones. Using that spectral signature information, the regression equation estimated a population density of 0.83 people per hectare and a total CBG population of over 431,000 inhabitants. It is an understatement to conclude that some manual adjustment of CBG boundaries to better fit the actual urban / suburban area within a CBG is warranted.

[21] Total population error was calculated as the absolute value of (true CBG population - modeled CBG population).

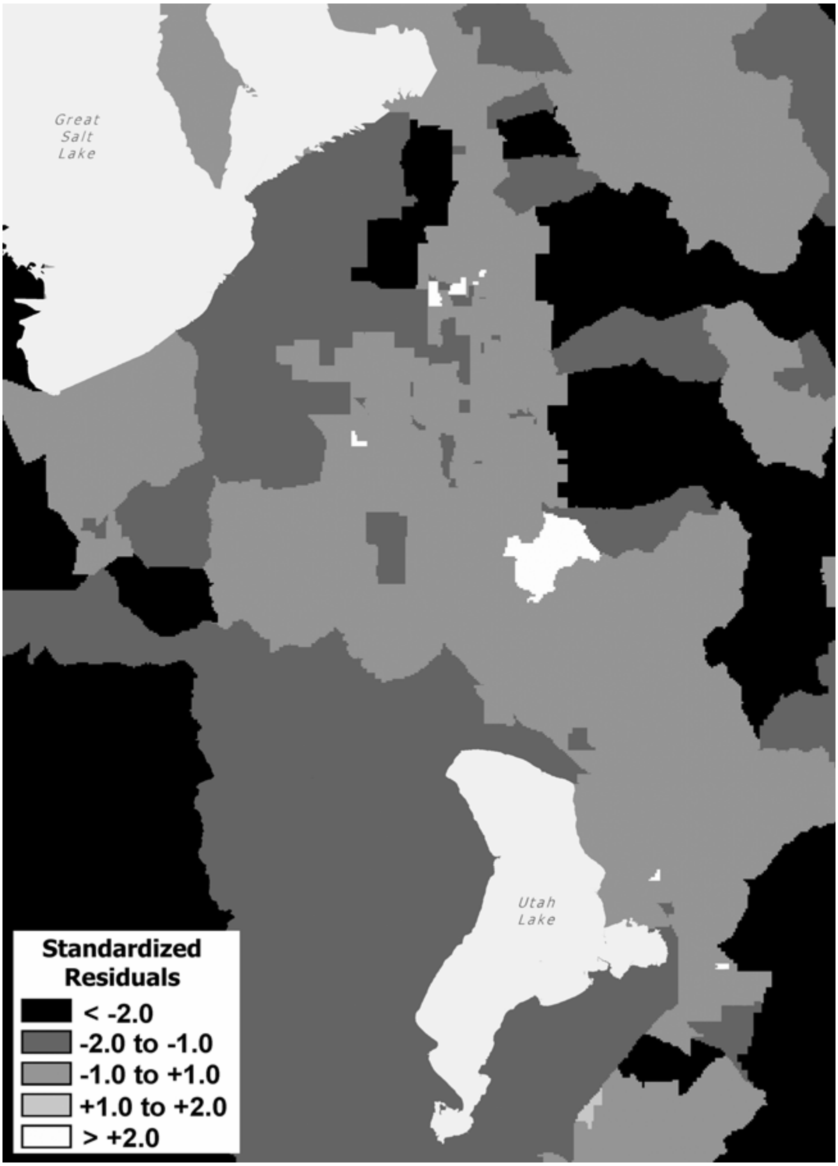

Figure 7. Map of residuals showing areas of over and underestimation by the regression equations

4.4 Concluding Comments

Research into population estimation using overhead imagery naturally leads to the question of whether the Decennial Census of the United States could be built upon remotely sensed data. In the early 1980's, there was a flurry of discussion regarding that question. The dialog was formally started by Brugioni (1983) who sided with the Secretary of the Commerce in his belief that the accuracy of the 1980 census could have been increased while concurrently decreasing the cost. At the time of Brugioni's writing, politicians in many large U.S. cities were unwilling to accept the enumeration as accurate, generally claiming the undercounting of urban subgroups. The primary issue was not apportioned representation in congress but rather the disbursement of federal funds based on population. City mayors observed that minorities and illegal immigrants used federally funded services and an accurate count of them was essential. Brugioni proposed, "It is time to stop and assess the prospects for better performance of the 1990 census. By using overhead reconnaissance systems which carry sophisticated cameras and remote sensing equipment, and by employing modern interpretation methods backed up by the latest computer technology, I am convinced that the census can be done more accurately, cheaper, faster, and better than by previous methods" (p. 1337). Given that the article was a commentary designed to elicit comments from the journal readership, Brugioni can be forgiven for such enthusiasm. The stated reasons for his certainty are hazy. Brugioni discusses the wealth of technology available for the task and recites a paragraph of remote sensing success stories such as weather prediction, forest inventory, and strategic intelligence. He then wonders why the Bureau of the Census has "not attempted to use these same technologies to determine the number of people living in a given area of a U.S. city" (p. 1338). Brugioni also cites the ongoing military use of overhead imagery to estimate urban populations abroad. Brugioni fails to mention that such foreign estimations seldom have an accuracy check. In conclusion, Brugioni appeals to his own authority, "From my nearly 40 years of experience in all phases of reconnaissance and analysis activity, I am thoroughly convince [sic] that a census using space-age technology is not only feasible but can be performed better and cheaper, and be more responsive to the needs of modern-day America."

Academics were quick to highlight the logical flaws of Brugioni's arguments as well as other problems with conducting a census with space-age technology. Morrow-Jones and Watkins (1984), both human / population geographers, believed that overhead imagery could never become the primary data source for the census. Several problems were articulated. They cited the practical hurdle of imaging the whole nation in a timely manner that could provide the same April 1st national snapshot as the census. Operational definitions (e.g., the distinction between *de jure* and *de facto* counting) long used by the census bureau and long valued by social scientists because of their constancy through several decades would change, "interrupting an exceptionally valuable source of evenly spaced historical data" (p. 230). While some census variables (e.g. house size) might be amenable to collection on overhead imagery, "their meanings would be changed to the detriment of long term comparative research" (p. 230). Morrow-Jones and Watkins also claimed that too little was known about using image characteristics as surrogates for social characteristics measured in the census. "Can these methods tell us the change in age structure, household composition, family income, race, sex ratio, or other *characteristics* of the people? This is a crucial part of the census and the largest drawback to the suggested method for improving it" (p. 231). Ethical considerations of privacy and image data use also troubled Morrow-Jones and Watkins. Sinclair (1984) was similarly troubled and wrote, "In America, is there not a right to ignore the Census and the Census taker? Or is it so important that we be counted that our own spy networks must be trained on us?" (p. 80). In conclusion, Morrow-Jones and Watkins admitted that a decennial census might be possible using remotely sensed data, but "the tradeoff would mean a great deal less information" (p. 232).

We suspect that Paul (1984) probably states the present convergence of opinion on this matter. First is the opinion that the use of remotely sensed data for human disciplines such as social science and public health has been historically undervalued and inadequately studied. The studies collected by Liverman et al. (1988) demonstrated the kind of progress that might be possible. However, as regarding population geography, while "remote sensing technology can be useful *in conjunction* with traditional demographic enumeration techniques [it] cannot be used as a replacement" (Paul 1984, p.1611). Expanding Paul's thought, we likewise do not consider remote sensing a replacement for traditional Census enumeration, but as an important source of urban data nonetheless. As reviewed by Lo (2006), many researchers have found that remotely sensed data is the only source of population information available in many developing countries. The intelligent use of imagery to augment traditional approaches may likewise be an efficient and accurate way of completing intercensal population estimates in areas of rapid urban growth. Furthermore, as scientists seek ways to improve the human condition, it may be profitable to consider *gathering new kinds of urban data* via remote sensing rather than developing new high-tech ways to capture traditional measures. See Jensen and Cowan (1999) for a review. For example, Weber and Hirsch (1992) demonstrated the calculation of spatial urban quality of life indices from imagery that are not amenable to a ground survey. Lo and Faber (1997) likewise demonstrated the generation of quality of life variables from overhead imagery. Unfortunately, the works of Weber and Hirsch as well as Lo and Faber have not been widely studied and have certainly not penetrated mainstream urban social science. Thus, although remote sensing will not likely supplant the typical Census enumeration, it nonetheless may provide critical measures of import to social scientists and policy makers in the pursuit of both knowledge and social justice.

References

Adeniyi, P.O. 1976. Applications of Aerial Photography to the Estimation of the Characteristics of Residential Buildings. *The Nigerian Geographical Journal*, 18:189-200.

Adeniyi, P.O. 1980. Land Use Change Analysis Using Sequential Aerial Photographs and Computer Techniques. *Photogrammetric Engineering and Remote Sensing*, 46:1447-1464.

Adeniyi, P.O. 1983. An Aerial Photographic Method for Estimating Urban Population. *Photogrammetric Engineering and Remote Sensing,* 49: 545-560.

Binsell, R. 1967. *Dwelling Unit Estimation from Aerial Photography.* Northwestern University Department of Geography.

Brugioni, 1983. The Census: It Can Be Done More Accurately With Space-Age Technology. *Photogrammetric Engineering and Remote Sensing*, 49:1337-1339.

Collins, W.G. and A.H.A. El-Beik. 1971. Population Census with the Aid of Aerial Photographs: An Experiment in the City of Leeds. *Photogrammetric Record,* 7:16-26.

Deering, D.W., J.W. Rouse, R.H. Haas, and J.A. Schell. 1975. Measuring Forage Production of Grazing Units from Landsat MSS Data. *Proceedings of the Tenth International Symposium on Remote Sensing of Environment.* Ann Arbor, MI: Environmental Research Institute of Michigan. Vol. 2. pp. 1169-1178.

El-Beik, A.H.A. 1967. *Air photo-interpretation applied to the study of urban land use and the determination of population densities.* Ph.D. thesis, University of Leeds, 214 pp.

GOPB (Governor's Office of Planning and Budget). 2005. *Economic Report to the Governor*. Salt Lake City, Utah: Government Printing Office. Available online at http://www.governor.utah.gov/dea. Last accessed October 15, 2006.

Green, N.A. 1956. Aerial Photographic Analysis of Residential Neighborhoods: An Evaluation of Data Accuracy. *Social Forces*, 35: 142-147.

Green, N.A. 1957. Aerial Photographic Interpretation and the Social Structure of the City. *Photogrammetric Engineering* 23:89-99.

Hadfield, S.M. 1963. *Evaluation of Land Use and Dwelling Unit Data Derived from Aerial Photography*. Chicago: Urban Research Section, Chicago Area Transportation Study.

Harvey, J.T. 1999. Estimation of Population Using Satellite Imagery. PhD Thesis, University of Ballarat.

Harvey, J.T. 2002a. Estimating Census District Populations from Satellite Imagery: Some Approaches and Limitations. *International Journal of Remote Sensing*, 23:2071-2095.

Harvey, J.T. 2002b. Population Estimation Models Based on Individual TM Pixels. *Photogrammetric Engineering and Remote Sensing,* 68:1181-1192.

Huete, A.R. 1988. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sensing of Environment,* 25:295-309.

Hung, Ming-Chih. 2002. Urban Land Cover Analysis from Satellite Images, *Proceedings of the Pecora 15 Conference*, Nov. 10-15, 2002, Denver, CO, USA. Available online at http://www.isprs.org/commission1/proceedings02/paper/00099.pdf . Last Accessed September 1, 2006.

Iisaka, J. and E. Hegedus. 1982. Population Estimation from Landsat Imagery. *Remote Sensing of Environment*, 12:259-272.

Jensen, J.R. 2005. *Introductory Digital Image Processing: A Remote Sensing Perspective.* 3rd Ed. Upper Saddle River, N.J.: Prentice-Hall. 526 p.

Jensen, J.R. 2007. *Remote Sensing of the Environment: An Earth Resource Perspective.* 2nd Ed. Upper Saddle River, N.J.: Prentice-Hall. 592 p.

Jensen, J.R. and Cowan, D.R. 1999. Remote Sensing of Urban / Suburban Infrastructure and Socio-Economic Attributes. *Photogrammetric Engineering and Remote Sensing,* 65:611-622.

Kraus, S.P., L.W. Senger, and J.M. Ryerson. 1974. Estimating Population from Photographically Determined Residential Land Use Types. *Remote Sensing of Environment*, 3:35-42.

Langford, M., D.J. Maguire, and D.J. Unwin. 1991. The Areal Interpolation Problem: Estimating Population Using Remote Sensing in a GIS Framework. Pages 55-77 in *Handling Geographical Information: Methodology and Potential Applications*. I. Masser and M. Blakemore (Eds.). New York: Longman Scientific and Technical.

Li, G. and Q. Weng. 2005. Using Landsat ETM+ Imagery to Measure Population Density in Indianapolis, Indiana, USA. *Photogrammetric Engineering and Remote Sensing*, 71:947-958.

Lindgren, D.T. 1971. Dwelling unit estimation with color-IR photos. *Photogrammetric Engineering and Remote Sensing*, 37:373-377.

Liverman, D., E.F. Moran, R.R. Rindfuss, and PC. Stern. 1988. *People and Pixels: Linking Remote Sensing and Social Science*. Washington, DC: National Academy Press. 256 pp.

Lo, C.P. 1995. Automated Population and Dwelling Unit Estimation from High-Resolution Satellite Images: a GIS Approach. *International Journal of Remote Sensing*, 16:17-34.

Lo, C.P. 1986. Accuracy of population estimation from medium-scale aerial photography. *Photogrammetric Engineering and Remote Sensing*, 52 (12): 1859-1869

Lo, C.P. 2006. Estimating Population and Census Data. *Remote Sensing of Human Settlements*. Manual of Remote Sensing. 3rd Ed. Vol. 5. M.K. Ridd and J.D. Hipple (Eds.). Falls Church, VA: American Society for Photogrammetry and Remote Sensing, pp. 337-377.

Lo, C.P. and B.J. Faber. 1997. Integration of Landsat Thematic Mapper and Census Data for Quality of Life Assessment. *Remote Sensing of Environment*, 62:143-157.

Lu, D. and Q. Weng. 2004. Spectral Mixture Analysis of the Urban Landscape in Indianapolis with Landsat ETM+ Imagery. *Photogrammetric Engineering and Remote Sensing,* 70:1053-1062.

Morrow-Jones, H. and J.F. Watkins. 1984. Remote Sensing Technology and the U.S. Census. *Photogrammetric Engineering and Remote Sensing*, 50:229-232.

Olorunfemi, J.F. 1984. Land Use and Population: A Linking Model. *Photogrammetric Engineering and Remote Sensing*, 50:221-227.

Paul, C.K. 1984. More on the Census. *Photogrammetric Engineering and Remote Sensing,* 50: 1610-1612.

Plane, David A. and Peter A. Rogerson. 1994. *The Geographical Analysis of Population*. New York: John Wiley & Sons, Inc.

Ridd, M.K. 1995. Exploring a V-I-S (Vegetation - Impervious surface – Soil) Model for Urban Ecosystems Analysis through Remote Sensing: Comparative Anatomy for Cities. *International Journal of Remote Sensing,* 16:2165-2185.

Sinclair, D. 1984. The Census: It Can be Done More Accurately with Space-Age Technology." *Photogrammetric Engineering and Remote Sensing,* 50:80.

Singh, A. 1989. Digital change detection techniques using remotely sensed data, *International Journal of Remote Sensing*, 10:989-1003.

UNCHS 2001. *The State of the World's Cities*. Nairobi, Kenya: United Nations Centre for Human Settlements.

Watkins, J.F. 1984. The Effect of Residential Structure Variation on Dwelling Unit Enumeration from Aerial Photographs. *Photogrammetric Engineering and Remote Sensing,* 50:1599-1607.

Weber, C. and J. Hirsch. 1992. Some Urban Measurements from SPOT Data: Urban Life Quality Indices. *International Journal of Remote Sensing*, 13:3251-3261.

Webster, C.J. 1996. The potential of urban texture measures in monitoring urbanization form space. *GIS in Asia,* T. Fung, P.C. Lai, H. Lin and A.G.O. Yeh, eds. Singapore: GIS Asia Pacific, pp. 309-321.

Weng, Q., D. Lu, and J. Schubring. 2004. Estimation of Land Surface Temperature-Vegetation Abundance Relationships for Urban Heat Island Studies. *Remote Sensing of Environment*, 89:467-483.