Assessing the Utility of Satellite Imagery with Differing Spatial Resolutions for Deriving Proxy Measures of Slum Presence in Accra, Ghana

Justin Stoler,[1] Dean Daniels, John R. Weeks, Douglas A. Stow, and Lloyd L. Coulter
Department of Geography, San Diego State University, San Diego, California 92182-4493

Brian Karl Finch
Department of Sociology, San Diego State University, San Diego, California 92182-4423

Abstract: Little research has been conducted on how differing spatial resolutions or classification techniques affect image-driven identification and categorization of slum neighborhoods in developing nations. This study assesses the correlation between satellite-derived land cover and census-derived socioeconomic variables in Accra, Ghana to determine whether the relationship between these variables is altered with a change in spatial resolution or scale. ASTER and Landsat TM satellite images are each used to classify land cover using spectral mixture analysis (SMA), and land cover proportions are summarized across Enumeration Areas in Accra and compared to socioeconomic data for the same areas. Correlation and regression analyses compare the SMA results with a Slum Index created from various socioeconomic data taken from the Census of Ghana, as well as to data derived from a "hard" per-pixel classification of a 2.4 m Quickbird image. Results show that the vegetation fraction is significantly correlated with the Slum Index (Pearson's r ranges from –0.33 to –0.51, depending on which image-derived product is compared), and the use of a spatial error model improves results (multivariate model pseudo-R^2 ranges from 0.37 to 0.40 by image product). We also find that SMA products derived from ASTER are a sufficient substitute for classification products derived from higher spatial resolution QB data when using land cover fractions as a proxy for slum presence, suggesting that SMA might be more cost-effective for deriving land cover fractions than the use of high-resolution imagery for this type of demographic analysis.

INTRODUCTION

Acquiring accurate census data in developing countries is often a difficult and time-consuming process. Lack of infrastructure and the increasing presence of dense, slum-like housing make collecting such data labor intensive and expensive at the neighborhood level (Baudot, 2001). The lack of reliable demographic data inhibits

[1]Corresponding author; email: stoler@mail.sdsu.edu

GIScience & Remote Sensing, 2012, **49**, No. 1, p. 31–52. http://dx.doi.org/10.2747/1548-1603.49.1.31

both urban planning and disaster recovery. So-called *data dissonance*, or the lack of coordination between existing demographic and geospatial data sets and their inaccessibility for responders, can result from data that are outdated, incorrectly scaled, or otherwise inappropriate for decision makers (National Research Council, 2007). The use of remote sensing imagery to collect or validate demographic characteristics has not been extensively examined. Many studies have utilized remotely sensed data to calculate population and map land cover and land use distributions (Ward et al., 2000; Chen, 2002; Lu and Weng, 2004; Wu and Murray, 2005), but there is little research on the correlation between demographic data and information extracted from remote sensing data. Such predictors can provide insight into the socioeconomic and health-related characteristics present on the ground (Rindfuss and Stern, 1998).

Although demographic variables cannot be accurately determined directly from remotely sensed imagery, the spatial characteristics of land cover elements such as rooftops, soil, and vegetation can be quantified to serve as proxies for the identification of both slum-like areas and residential areas associated with higher socioeconomic status (Weeks et al., 2007). These socioeconomic characteristics of urban areas could provide a framework for analysis of health inequalities and areas of poverty so that researchers and policymakers are able to make decisions about which areas are most in need of attention.

Spatial resolution is one factor to consider when using imagery to infer demographic characteristics from land cover. Moderate-spatial-resolution imagery such as ASTER and Landsat-5 TM are economically priced or free, more accessible, and typically offer wider spatial coverage than more expensive high spatial resolution imagery. But differences in spatial resolution can result in differing representations of the land cover in spatially heterogeneous regions and may thus significantly impact land cover classification. Due to the highly complex and heterogeneous spatial structure found in Third World urban cities, pixel-level classification techniques may not enable detection of small changes in the land cover, especially when using imagery of moderate spatial resolution. Given this concern, it is important to assess the accuracy of sub-pixel classification techniques based on moderate spatial resolution imagery relative to the "hard" per-pixel based methods from higher spatial resolution data when classifying land cover in developing cities.

The objective of this paper is to assess the utility of satellite remote sensing data as an indicator of urban slum conditions in Accra, Ghana, a city typical of developing urban areas. We focus on three key sets of questions. (1) How does choice of satellite imagery affect the reliability of land cover fraction estimates, and how appropriate is moderate- (i.e., economical) spatial-resolution imagery for demographic analysis? (2) Will a sub-pixel technique (spectral mixture analysis) improve relationships between land cover and socioeconomic factors relative to a per-pixel-based classifier, and which technique quantifies land cover in a manner that is most closely associated with measures of land use and socioeconomic status derived from census data? (3) Are percentages of land cover and socioeconomic variables correlated, and if so which types of land cover provide the strongest correlation? The paper concludes with a discussion of the implications of spatial resolution, classification technique, and land cover categorization for producing new types of demographic variables for analysis in areas where such data are scarce or unreliable.

BACKGROUND

Heterogeneity and Scale in Urban Demographic Analysis

Heterogeneity in the urban environment can be a significant hurdle for using satellite imagery to classify land cover given the diversity of built structures, vegetation types, bare soil zones, and water bodies (Herold et al., 2002). The theoretical construct for the use of remotely sensed data is based on the idea that human behavior is shaped by these natural, built, and social environments in which they reside (Rashed et al., 2005). The use of remote sensing imagery to analyze demographic patterns has its roots in urban ecosystem classification (Anderson et al., 1976). The type of classification process employed is important when attempting to infer demographic information from the imagery. "Hard" classification, which allows only one class type per pixel, commonly assigns classes to a pixel based on the highest probability (or other algorithmic determination) of membership (Zhang and Foody, 1998). This type of classification can result in misrepresentation or inaccuracies over land cover that is highly heterogeneous, as pixels are typically a representation of a mixture of surface types that can create confusion for classification algorithms (Cushnie, 1987). This "mixed pixel problem" is exacerbated when imagery is too coarse and pixels are larger than features on the ground, but the utilization of increasingly higher spatial resolution data may also inhibit classification accuracy as a result of similar within-class spectral variability, which can cause confusion of feature class assignments (Cushnie, 1987; Marceau, 1999). Fisher (1997) describes four main causes of this mixed pixel problem: (1) boundaries between two or more mapping units; (2) the transition between types of mapped phenomena; (3) linear sub-pixel objects like a road; and (4) small sub-pixel objects like a house or tree. All of these challenges tend to be common in urban remote sensing applications, and higher spatial resolution is ultimately necessary in urban morphological analysis due to the heterogeneity of urban structure and the similarity between certain spectral responses (Herold et al., 2002). Previous investigations of the tradeoffs between the accuracy and cost of moderate versus high spatial resolution imagery (e.g. Johansen et al., 2010) underscore the importance of the effects of scale.

Research on the effects of scale continues to highlight the modifiable areal unit problem (MAUP) (Openshaw, 1984), for which remote sensing may be viewed as a particular case (Marceau and Hay, 1999). An analysis of the effects of scale using three sets of satellite imagery with ground sampling distances (GSD) ranging from 20 m to 1.1 km found that most of the variables were sensitive to changes in spatial resolution, but with no real consistent pattern (Benson and Mackenzie, 1995). Remote sensing applications ranging from forestry classification (Marceau et al., 1994) to population density estimation (Wu and Murray, 2005) have found that a correlation between variables may exist at one scale but not another, which potentially creates biased results. Others continue to acknowledge that future research is needed to effectively assess the impact of the aggregation of metric data resulting from a change in scale (Herold et al., 2002). Consequently, conducting analyses with varying spatial resolutions can be an effective way to study potential MAUP effects on the sub-pixel land cover classification process as well as the collection of the demographic data of interest to this study.

Spectral Mixture Analysis

In cities of developing countries, the level of material composition can be so complex that assigning one class to each pixel can yield an inaccurate representation of land cover when using moderate spatial resolution imagery (Chen, 2002). A sub-pixel classification may provide a more realistic representation of land cover in Accra by accounting for such heterogeneity. Spectral mixture analysis (SMA) estimates the proportion of dominant material elements that compose the ground resolution element associated with a pixel by quantifying the percentage of particular land cover materials from a number of determined endmembers (Iverson et al. 1989). These endmembers are a class of spectrally pure pixels that represent known land surface materials and are used as training-site inputs for the SMA algorithm. The percentages of each class within a pixel are calculated based on the level of certainty that a given pixel belongs to a particular class (Cross et al., 1991). The algorithm also generates fraction images for each endmember, which represent the observed mixed pixel reflectance spectrum (Van der Meer and De Jong, 2000).

A widely utilized method for analyzing urban spatial patterns and identifying candidates for endmember selection is the VIS model, which assumes that land cover in urban environments is a linear combination of three components: vegetation, impervious surface, and soil (Ridd, 1995). These three land cover types are modeled as the fundamental components of the urban environment, and this model has demonstrated improved classification accuracy using the SMA technique on Landsat Thematic Mapper (TM) imagery (Ward et al., 2000; Phinn et al., 2002), Landsat Enhanced Thematic Mapper Plus (ETM+) imagery (Wu and Murray, 2003; Lu and Weng, 2004), and multi-spectral imagery (Rashed et al., 2001, 2005).

VIS does not account for land cover types such as water and wetlands, both prominent in Accra's landscape, and thus requires masking of such features prior to image classification. Confusion between types of impervious surface can also exist; classification errors often result from confusion between bare rocks, vehicle tracks, land under construction, or otherwise complex surface material composition (Chen, 2002; Lu and Weng, 2004), which could be problematic given Accra's large proportions of dirt roads and razed land. Higher spatial resolution imagery may minimize this problem by maximizing the number of pixels which make up an urban area, thereby creating a more detailed representation of the landscape. Despite these challenges, the VIS model still remains the best available framework for delineating endmember candidates for an SMA framework.

Prior research has demonstrated links between the urban landscape and demographic patterns. Accra's urban structure was explicitly investigated in a study of slum presence and neighborhood characteristics that identified a distinct correlation between land cover and socioeconomic status at the neighborhood level (Weeks et al., 2007). Ridd's VIS model was used as the basis for a per-pixel–based classification of high spatial resolution QuickBird imagery (2.4 m) in order to create land cover percentage data for each neighborhood. These percentages were combined with various socioeconomic data to produce a regression model for predicting slum presence, and the proportional abundance of vegetation is a significant predictor of slum status. Stoler et al. (2009) noted the spatial structure of socioeconomic status relative to urban agricultural activity in Accra. These relationships between socioeconomic status and

land cover in Accra affirm the potential utility of satellite-derived data to infer additional demographic characteristics.

METHODOLOGY

Study Area and Data

Ghana's capital city, Accra, is located on West Africa's south-facing coast along the Gulf of Guinea. Accra's population exceeded 1.6 million in the 2000 census, and is estimated to have reached over 2.3 million in 2010 while still growing over three percent annually (United Nations, 2010). Accra has been the seat of Ghana's government since 1887 and became the capital in 1957 when Ghana became the first sub-Saharan country to gain independence. For administrative purposes, Ghana Statistical Service has partitioned the Accra Metropolitan Area (AMA) into over 1,700 Enumeration Areas (EAs), which are roughly equivalent to U.S. census block groups or U.K. output areas. Figure 1 portrays Accra's location in Ghana and sub-Saharan Africa, as well as the EA distribution within Accra.

The urban landscape of the greater Accra region is heavily influenced by its colonial history. A distinct neighborhood structure has evolved from the *de facto* segregation and social exclusivity brought on by European occupation in the 19th century (Pellow, 2002). This spatial grouping by class and ethnicity developed into an unwritten part of the city's constitution and allowed for unequal access to resources and distribution of facilities, a trend that more recently has been exacerbated by rapid population growth coupled with local government favoritism and corruption (ibid.). Consequently, studies have found there is a strong presence of societal and health-related inequalities that are spatially distributed in a complex heterogeneous manner (Weeks et al., 2006, 2007). Likewise, studies have found a great deal of spatial autocorrelation at the Enumeration Area level, which indicates a distinct spatial distribution with slum presence (Weeks et al., 2007) and other socioeconomic measures (Weeks et al., 2010).

Two moderate spatial resolution and one high spatial resolution multi-spectral satellite images are utilized in this study; the image characteristics are summarized in Table 1. An ASTER image from April 27, 2001 contains visible and near infrared (VNIR) bands (spectral range of 0.52–0.86 μm) that have a spatial resolution of 15 m, and cover a swath width of 60km. The shortwave infrared (SWIR) and thermal infrared wavebands are not utilized because their spatial resolutions (60 m and 90 m, respectively) are coarser than the VNIR bands.

Another moderate spatial resolution image utilized for this study is a Landsat 5 Thematic Mapper (TM) image acquired February 4, 2000 with 30 m spatial resolution and a 185 km swath width. Bands 3 (red, 0.63–0.69 μm), 4 (near infrared, 0.75–0.90 μm), and 5 (shortwave infrared, 1.55–1.75 μm) are used from this image because band 2 (green, 0.525–0.605 μm) suffered from illumination distortion. Band 2 was subjected to a cross-track illumination correction tool in an attempt to reduce the effects of this illumination distortion, but most of the image covering Accra remained affected. The shortwave infrared band is used in place of the green band as it yields the best contrast and detail within urban areas of the remaining bands. TM Band 7 (SWIR2) is

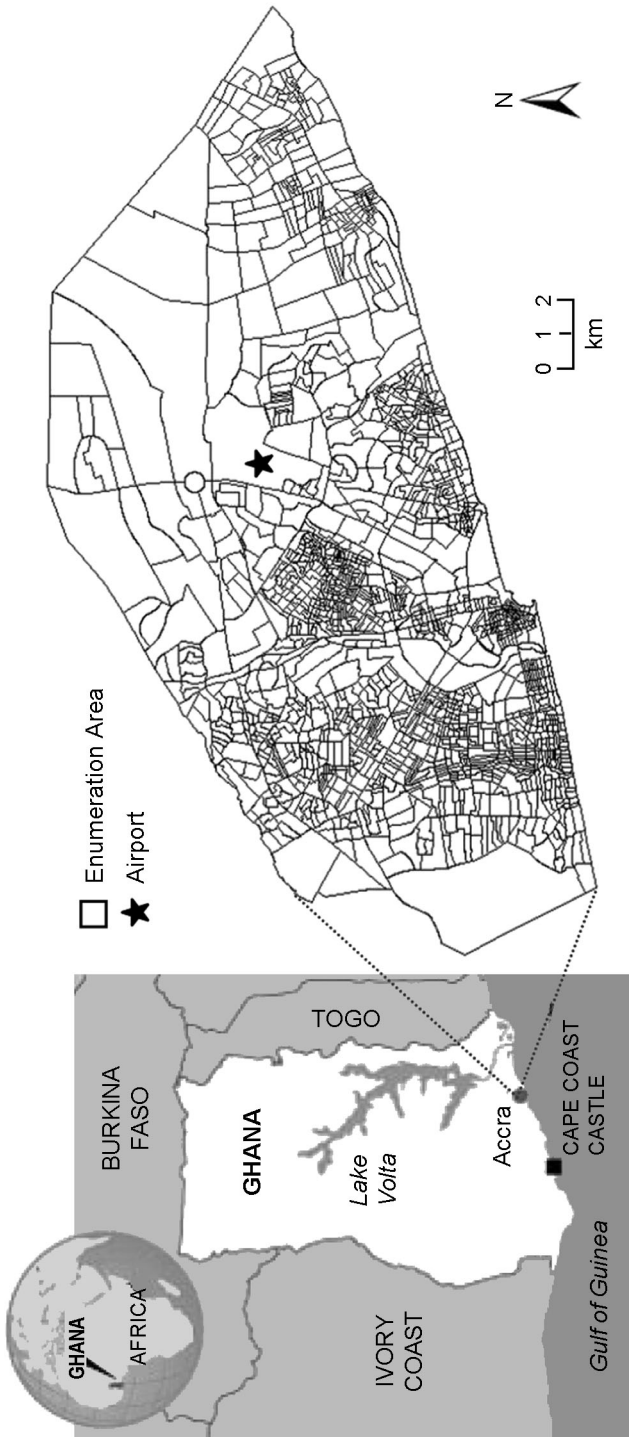

Fig. 1. Study area in Accra, Ghana, with GIS layer of Enumeration Areas.

Table 1. Image Characteristics of QuickBird, ASTER, and Landsat TM Scenes Analyzed

	QuickBird	ASTER	Landsat TM
Acquisition date	April 4, 2002	April 27, 2001	February 4, 2000
Swath	16.5 km	60 km	185 km
Spatial resolution	2.4 m	15 m	30 m
Wavebands used	2, 0.52–0.60 μm	2, 0.52–0.60 μm	3, 0.63–0.69 μm
	3, 0.63–0.69 μm	3, 0.63–0.69 μm	4, 0.75–0.90 μm
	4, 0.76–0.90 μm	4, 0.76–0.86 μm	5, 1.55–1.75 μm

not used as an input to the SMA models, as it mostly provides redundant information to Band 5.

The third and highest spatial resolution image used is a 2.4 m QuickBird 2 (QB) multispectral image acquired April 4, 2002, with spectral coverage ranging from visible to near infrared (0.45–0.90 μm) and spatial extent of 16.5 km by 16.5 km. The spatial extent of coverage of the QuickBird image does not encompass all of the Accra Metropolitan Area. A substantial part of this analysis involves correlating SMA results to the data derived from this image, and therefore the extent for these statistical analyses and comparisons were limited to that of the QuickBird image. Thus 237 of 1,724 Enumeration Areas are excluded from the analysis, with most of these located in the eastern region of the city.

Our retrospective approach is limited by the available image data that had been captured during three different years (2000, 2001, and 2002), as even tasked image acquisitions for tropical locations such as Accra are beset by perpetual cloud cover. The three acquisition dates occur seasonally within the onset of Accra's rainy season, which generally ramps up in April and is wettest in May and June. We were unable to locate adequate daily or monthly rainfall data for Accra to determine inter-annual differences from 2000–2002, but there is a potential bias toward greater vegetation cover in the ASTER image acquired April 27, 2001 relative to the Landsat TM (February 4, 2000) and Quickbird images (April 4, 2002).

A geographically referenced GIS layer of Accra's enumeration areas in shapefile format (see Fig. 1) is used to extract the average land cover fraction composition for each EA within the study area. This layer was previously created by manual digitization of EA boundaries based on descriptions provided by Ghana Statistical Service and interpretation of the georeferenced QuickBird imagery.

Approach

The first objective of this study is to generate land cover fraction data through image classification and spectral mixture analysis (SMA), and then compare fraction estimates from various data sources and approaches at the level of Enumeration Areas. SMA is performed on both the ASTER and Landsat TM images and the SMA results are compared to VIS percentages created from per-pixel hard classification of the ASTER and QuickBird image. The second objective is to determine how correlated

SMA land cover fractions are with Slum Index data from Accra (a socioeconomic index described below), and which land cover types and image-derived products provided the best indicator of slum characteristics. If the SMA data are highly correlated with variables derived from socioeconomic data (such as the Slum Index), then such an approach could be applied to generate supplementary data for other Third World cities that lack reliable or abundant demographic information.

Image Pre-processing

Several pre-processing steps are taken to prepare the imagery for information extraction and analysis. Because the QB image is used as a source for delineating endmembers—the higher spatial resolution helps to identify the land cover type corresponding to pixels determined to be spectrally pure—the ASTER and TM images are spatially registered to the QB image (root mean square error [RMSE] < 0.5 pixels using 20–25 ground control points) in order to compare land cover proportions derived from the image sets. Water features are masked in the analysis because they can be problematic when attempting to quantify land cover fractions using SMA, because directional effects of reflected sunlight can result in classification confusion with urban surfaces. Further, a water endmember is not a component of the desired VIS classes. The masks are manually digitized along the borders of water bodies. Once the images are registered and masked, endmember selection for SMA commences. All processing of remotely sensed data is performed using ENVI, a software application created by ITT Visual Information Solutions.

Per-pixel Classifications

Per-pixel unsupervised classification of VIS components is performed on both the QB and ASTER images (only SMA was performed on the Landsat image due to its coarser spatial resolution). Fifty spectral cluster classes are specified for each image set, and pixels are classified into one of these cluster classes. The cluster classes are subsequently labeled as vegetation, impervious surface, or soil land cover classes based on visual interpretation of the image and georeferenced field data collected in Accra. Only a single cluster from the ASTER classification—corresponding to both known cleared land and large buildings—requires further classification (i.e., "cluster busting") into 10 additional clusters, yielding more precise VIS class assignments. This procedure is unnecessary for the QB-derived unsupervised classification product.

SMA Estimates of Land Cover Fractions

Endmember (EM) selection is the crucial first step in the spectral unmixing process that can be achieved through SMA. The goal is to determine the most extreme (in spectral-radiometric feature space) or spectrally pure pixels for inherent land cover (LC) classes as the primary signature bases for unmixing algorithms. EMs are extracted by manually selecting pixels of interest from the spatial, statistical, or feature space representations of the image. This image retrieval method of selecting EMs (compared to spectral library methods) is more appropriate for this analysis due to available ancillary data, the lack of atmospheric optical data for atmospheric correction, and the

absence of known spectral libraries for the study area. Studies which select EMs using the image retrieval method often use some form of high spatial resolution imagery to identify features on the ground even when the spectral unmixing is being performed on coarser imagery (Lu and Weng, 2004; Rashed et al., 2001; Ward et al., 2000). Because part of the rationale of this analysis is to assess the practicality of using moderate spatial resolution imagery such as ASTER and Landsat in place of high spatial resolution imagery, EMs are selected both with and without using the QB image in order to assess the need for high spatial resolution imagery for SMA EM selection.

Vegetation, impervious surface, and soil (VIS) are the three endmember classes of interest. A water or shade class is unnecessary due to the masking of water features, the presence of relatively few high-rise buildings that can cast shadows, and the equatorial location of Accra with small solar zenith angles (and therefore minimal shadowing). A set of endmember candidates is initially selected without the use of the QB imagery, based solely on the Pixel Purity Index (PPI) and visual analysis of feature space scatterplots. The PPI is an iterative process that repeatedly projects n-D scatterplots on a random unit vector and records the pixels that fall on the ends of the unit vector. The pure pixels are assigned a value based on the number of times the index recorded the pixel as extreme, with higher values representing pixels that are more spectrally pure. The PPI creates an output image of the pure pixels with values, which can then be overlaid onto the original image for analysis.

Once spectrally pure pixels are located using the PPI, the corresponding locations on the image are determined using 2-D scatterplots. Ideal endmember candidates are often located at the tips or inflections of these plots (Rashed et al., 2001). Vegetation and soil EMs are most discernible, while pixels representing pure impervious surfaces are usually more difficult to identify along the bottom axis of the tasseled-cap feature space distribution due to the diverse types of impervious materials sensed in the urban landscape of Accra. The Accra airport (Fig. 1) is ultimately used to identify pure impervious surface pixels by examining its large buildings and wide, spectrally homogeneous runway.

The combination of high spatial resolution imagery and scatterplot analysis is an effective method of selecting endmember signatures (Lu and Weng, 2004). Using similar procedures, we select a second set of endmember candidates by incorporating visual analysis of the high spatial resolution QB imagery in addition to the tools described above.

We use an iterative approach for both methods of EM extraction. This involves locating potential EM candidates, running the SMA model, then assessing the output and spatial characteristics of the RMSE. Purer endmembers are then selected based on RMSE and presence of negative and "super positive" endmember values. Super positive endmembers are fraction values that exceed 100%, which constitute pixels that are potentially purer than those used in the unmixing process. Negative fraction values from one land cover type usually correspond to superpositve fractions from a different endmember type. The SMA model is run again with the purer endmembers selected from these superpositve regions. This process is repeated until an output is reached that minimizes these superpositive and negative values. Both sets of endmembers (selected with and without guidance from the QB imagery) are run through the SMA model to assess whether the added cost of QB for EM selection is warranted. The spectral unmixing process is conducted using the SMA routine in ENVI with unconstrained

variables that allow for the inclusion of negative values and do not require fractions to sum to one. Fraction images are produced for each endmember (vegetation, impervious surface, and soil) for both the ASTER and Landsat TM data sets.

Once the fraction images are created, the average fraction of each land cover (VIS) class is derived for each EA using the EA boundary layer. Since the spectral unmixing process is unconstrained to allow for error, most of the values for each EA do not sum to 1. To control for this and create consistent percentages across data sets, the fraction values for each EA are normalized by dividing the value of each land cover type by the sum of all three within its EA.

Statistical Analyses

Pearson's correlation coefficient matrices are generated to assess the association between VIS fractions derived from different image types and image analysis approaches, and also to assess the correlation of each set of image-derived fraction data with a Slum Index generated from census data. The first set of correlation coefficients is used to examine the agreement between VIS estimates from SMA of ASTER and Landsat TM imagery, and from per-pixel classifications of ASTER and QB data.

Slum Index (SI) scores are calculated using a 10% random sample of individual-level housing characteristics collected by Ghana Statistical Service in the 2000 census for all 1,724 EAs in the study site (Weeks et al., 2007). The survey questions used to create the Slum Index are based on UN-Habitat criteria for a slum dwelling. Each housing unit in the 10% sample is allocated one point for each of the following criteria it met: (1) lack of piped water within the unit; (2) lack of toilet and sewage connection; (3) the number of persons per room exceeds two; (4) less durable building materials; (5) the resident is not the owner. The SI for each housing unit is the sum of these met criteria, so that 0 indicates a house with no slum characteristics and 5 indicates a house where all five criteria were met. SI scores for each EA are created by calculating the average slum score for all housing units within each of the 1,487 EAs completely covered by the extent of the QB imagery. Despite potential reporting biases in the 2000 census, the SI is a useful measure for identifying Accra's most and least slum-like neighborhoods (Weeks et al., 2007).

Univariate and multivariate regression analyses are also performed using the image-derived land cover fractions as independent variables on the SI scores summarized by EA. Due to the linear nature of the VIS model, a multiple regression model cannot explain the relationship with SI if all three classes are included. Thus, the correlation coefficients are used to determine the two land cover variables with the highest correlation to the SI, which are subsequently used in a multiple regression model. Residuals from the SI regression models are examined for outliers and influential values beyond a 95% confidence interval. Identified outliers are evaluated to highlight any land cover characteristics that may create confusion in the regression model.

The presence of any spatial trend in the data would violate the regression assumption of spatial independence (Rogerson, 2001). Thus, regression residuals are also tested for spatial autocorrelation using Moran's I statistic in order to detect any pattern of spatial dependence in the residuals. A positive Moran's I value indicates spatial clustering, while a negative value indicates a pattern of dispersal, and the associated

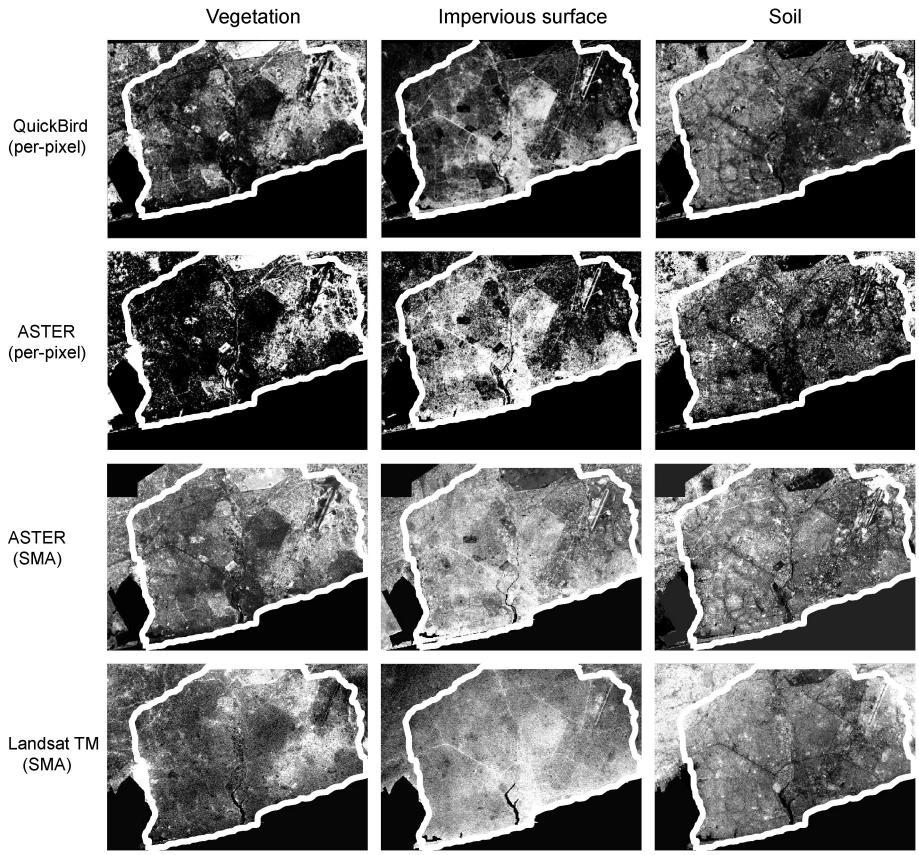

Fig. 2. Land cover fractions derived from per-pixel classification of QuickBird and ASTER, and SMA of ASTER and Landsat TM images, with common study area outlined.

Z-score allows a test of statistical significance for any spatial autocorrelation present. Once detected, spatial dependence is accounted for in the regression analysis by employing a spatial error model in the software program GeoDa v.0.9.5-I (Anselin et al., 2006). This model adds a spatial parameter to the regression equation in order to control for the spatial dependence in the data. The R^2 measure from the spatial error model is compared to the original regression output to evaluate the impact of spatial clustering on model prediction.

RESULTS AND DISCUSSION

SMA Results

The VIS fractions are portrayed in Figure 2. The PPI is effective at identifying pure vegetation pixels for the ASTER image, and EM candidates selected are generally similar to those selected based on interpretation of the QB image. The PPI does not perform well for soil or impervious endmembers, suggesting that high spatial resolution

imagery for a limited extent may be necessary when deriving these endmembers. Pure soil pixels for the ASTER image are difficult to identify with high certainty due to confusion with bright rooftops and spectral similarity with certain impervious materials, and we use the QB image to successfully locate areas of dry bare soil for EMs. Similarly, candidate pixels representing impervious surface EMs are difficult to identify based on analysis of red and NIR feature space displays. Impervious surfaces associated with the airport are distinct enough to be identified on the ASTER image, but the spatial resolution is not fine enough to distinguish the specific material composition of the surface (roof top, bright asphalt, etc.) identified by the PPI. Visual inspection of the QB image reveals that the PPI is only able to highlight a few of the rooftops and bright patches of asphalt around the airport buildings. Selecting candidate soil and impervious pixels associated with superpositive fractions from the initial SMA run help to alleviate this confusion.

The PPI, when utilized for selecting Landsat TM endmembers, is complicated by coarser spatial resolution and identifies many potentially pure pixels that are not located along the tips of the vertices in the 2-D data cloud and, upon visual inspection with the QB, are a mixture of two or three land cover types. Thus, the QB image is necessary to adjust the PPI output, particularly for soil and impervious selections. Despite this confusion there is some success identifying vegetation EMs using the PPI from the Landsat TM image. The PPI identifies a few pixels which resemble the ASTER vegetation EM candidates in spectral space, though many vegetated areas selected in the ASTER PPI process are not identified in the Landsat TM analysis, probably due to the coarser spatial resolution or the absence of the green band. The PPI identifies pure impervious pixels that actually represent a mixture of impervious surface and soil. As with ASTER, the QB image is used as a reference to produce a superpositive pixel from the airport area. The PPI also identifies pure areas of dry bare soil from Landsat TM with greater success after visual comparison with the QB image.

The first row of Table 2 provides mean land cover fraction values derived from each image source. A greening bias due to the later acquisition day of the ASTER image relative to the Landsat TM and Quickbird images appears to be evident. While the impervious surface fractions are relatively constant across data products, the vegetation fraction is consistent with the temporal pattern of average monthly precipitation in Accra, varying inversely with the soil fraction.

ASTER: SMA vs. Per-pixel Classification

Table 2 also contains the correlation coefficients of full-image VIS fractions and percentages from all image types and classification approaches for comparison. The ASTER per-pixel derived land cover fractions exhibit high to very high correlations with the QB percentages: the impervious surface class exhibits the strongest correlation ($r = 0.905$), followed by vegetation ($r = 0.884$) and soil ($r = 0.826$). Of the SMA fractions, only vegetation yields a stronger correlation with the QB percentages than the per-pixel percentages (0.972 vs. 0.884), suggesting that the SMA fractions offer better characterization of vegetation. The SMA impervious fractions are slightly less correlated with the QB percentages than were the per-pixel percentages (0.863 vs. 0.905), while the SMA soil fractions are substantially less so (0.577 vs. 0.826).

Table 2. Mean Land Cover Fraction Values (n = 1487 EAs) with Pearson Correlation Coefficients of QuickBird Per-pixel Land Cover Fractions, ASTER Per-pixel Fractions, SMA Fractions from ASTER and Landsat TM, and the Slum Index

Land cover fraction	Quickbird (per-pixel)			ASTER (per-pixel)			ASTER (SMA)			Landsat TM (SMA)			Slum Index
	Veg	Imp	Soil	Veg	Imp	Soil	Veg	Imp	Soil	Veg	Imp	Soil	
Mean fraction value	0.178	0.541	0.282	0.065	0.633	0.302	0.253	0.539	0.208	0.132	0.523	0.345	
QuickBird (per–pixel)													
Veg	1												–0.495**
Imp	–0.792**	1											0.450**
Soil	0.196**	–0.754**	1										–0.190**
ASTER (per–pixel)													
Veg	0.884**	–0.546**	–0.073**	1									–0.330**
Imp	–0.770**	0.905**	–0.626**	–0.605**	1								0.371**
Soil	0.404**	–0.785**	0.826**	0.125**	–0.866**	1							–0.254**
ASTER (SMA)													
Veg	0.972**	–0.815**	0.264**	0.833**	–0.777**	0.444**	1						–0.509**
Imp	–0.793**	0.863**	–0.534**	–0.573**	0.858**	–0.709**	–0.855**	1					0.406**
Soil	–0.132**	–0.271**	0.577**	–0.318**	–0.327**	0.607**	–0.060*	–0.466**	1				0.088**
Landsat TM (SMA)													
Veg	0.769**	–0.477**	–0.060*	0.735**	–0.498**	0.158**	0.775**	–0.518**	–0.324**	1			–0.385**
Imp	–0.679**	0.865**	–0.659**	–0.446**	0.786**	–0.699**	–0.743**	0.850**	–0.369**	–0.443**	1		0.345**
Soil	0.355**	–0.715**	0.766**	0.114**	–0.616**	0.696**	0.424**	–0.677**	0.581**	–0.028	–0.884**	1	–0.183**

** $p < 0.01$ (2–tailed); * $p < 0.05$ (2–tailed).

There are advantages and disadvantages associated with performing SMA on ASTER imagery as a substitute for QB imagery compared to performing a hard per-pixel classification of the ASTER data. The SMA-derived vegetation fraction demonstrates greater agreement with QB than the per-pixel percentage, while the impervious fraction shows little difference. Hence the viability of using SMA on ASTER data rather than higher spatial resolution imagery depends on which VIS classes are theorized to be linked to the process of interest.

SMA Fractions vs. QuickBird Per-pixel Classification

We emphasize the comparison of SMA fractions derived from moderate spatial resolution image data with QuickBird-derived estimates, since theoretically the latter should enable derivation of more accurate estimates and is often considered as reference data for assessing the accuracy of the former (Small and Lu, 2006). Most of the correlation coefficients for land cover fraction estimates between the SMA and QB data in Table 2 are significant at a 99% confidence level ($\alpha = 0.01$). Because statistical significance is influenced by the large sample size (1,487 Enumeration Areas), which can help detect smaller statistical effects (Hair et al., 1998), this analysis focuses primarily on the correlation coefficient (r) rather than the significance assigned to each relationship. In lieu of definitive guidelines for interpreting the strength of correlation statistics, we utilize the following framework based on Cohen (1982): 0 to 0.19 is very low, 0.20 to 0.39 is low, 0.40 to 0.69 is moderate, 0.70 to 0.89 is high, and 0.90+ is very high.

Vegetation fractions derived from SMA on ASTER data exhibit the strongest correlation to estimates from the per-pixel classification of QB data ($r = 0.972$) of any fraction estimate in this study, and is the only relationship considered *very high* on Cohen's scale. The SMA-derived ASTER impervious fraction is *highly* correlated with the QB impervious proportion ($r = 0.863$), especially when considering the complexity and variation in impervious surface type across Accra. The ASTER soil fraction exhibits a weaker correlation ($r = 0.577$) with the QB soil data, suggesting that the difference in spatial resolution may have an impact on estimating soil fractions. This could be attributed to confusion between spectra during the EM selection process, or to the lack of a representatively pure soil pixel at the 15 m scale of the ASTER image.

Table 2 also shows strong negative correlations between different fraction variables, which bear substantive significance. The ASTER impervious surface fractions are highly negatively correlated with the QB vegetation percentages ($r = –0.793$), as are the ASTER vegetation fractions with the QB impervious percentages ($r = –0.815$). These are nearly identical to the correlation between QB vegetation and QB impervious surface ($r = –0.792$). Substituting SMA-derived ASTER fractions for QB land cover percentages of either vegetation or impervious surfaces produces almost the same correlation values, which indicates further agreement between the two data sets.

The SMA-derived Landsat TM fractions also exhibit strong correlations with the QB data. Impervious surface exhibits the strongest relationship ($r = 0.865$) and is nearly identical to the ASTER impervious fraction correlation (0.863), despite the difference in spatial resolution. The correlation of the Landsat TM vegetation fraction with QB vegetation is considerably weaker ($r = 0.769$) than the ASTER vegetation fraction (0.972), whereas the Landsat TM soil fraction is substantially more correlated

(r = 0.766) with QB than the ASTER soil fraction (0.577). It is unclear whether these discrepancies are due to spectral confusion or an underlying component in either data set that is linked to vegetation or soil. Overall there is strong agreement between the QB- and SMA-derived data, and moving from 15 to 30 m spatial resolution yields just a slight weakening of the vegetation relationship while strengthening the soil relationship. These patterns are visible in scatterplots of the ASTER and Landsat TM fractions vs. the QB proportions (Fig. 3), which highlight the variation between the data sets.

Modeling Slum Characteristics

The last column of Table 2 contains the correlation coefficient statistics comparing each set of VIS land cover data to the SI. Although none of the correlations between each data set and the slum index are high, some of the fraction data exhibit moderate correlations.

Vegetation exhibits the strongest correlation with the SI for all data sets except for the per-pixel ASTER- derived data. The vegetation fraction from the ASTER SMA data has the strongest correlation (r = –0.509), higher than both the QB (r = –0.495) and TM SMA (r = –0.385) vegetation fractions. The per-pixel ASTER classification produces the lowest correlation of vegetation with the slum index (r = –0.330). These numbers indicate an inverse relationship between vegetation and the SI, meaning that more impoverished slum-like areas are generally associated with sparser vegetation. It is clear that the SMA-derived ASTER fractions have potential utility for predicting slum properties, as they display a stronger correlation with the SI than both the per-pixel ASTER fraction and the QB fraction. The Landsat TM vegetation data exhibit the weakest relationship with the SI, indicating that the 30 m spatial resolution may not be detailed enough to quantify the low levels of vegetation cover associated with slum conditions.

The impervious fractions also yield moderate positive correlations with the slum index, suggesting that greater amounts of impervious surfaces generally coincide with slum-like characteristics of a neighborhood. The QB and SMA ASTER impervious fractions produce similar correlations (r = 0.450 and 0.406, respectively), while the per-pixel ASTER impervious data are more weakly correlated with the SI (r = 0.371). This is curious when considering that the per-pixel ASTER impervious fraction is more highly correlated with the QB impervious fraction than with the SMA ASTER fraction. As with the vegetation fractions, the Landsat TM impervious fraction yields the weakest correlation (r = 0.345) relative to the other data sets.

Soil fractions produce the weakest correlations with the slum index of all land cover categories across all data sets. Soil does the poorest job of relating to the SI, not only by virtue of the weak correlations, but also because correlations were higher for the coarser data set for both the SMA and per-pixel methods of land cover fraction extraction. It is unclear what causes these discrepancies, but combined with the low r values, indicates that soil may not be suitable as a proxy for inferring "slumness" regardless of the spatial resolution or type of image processing method utilized.

Given the results from correlation analyses, vegetation and impervious surface are designated as independent variables for the multiple linear regression analysis. These fractions have stronger relationships with the SI than soil for both the ASTER and TM SMA, and both exhibit strong correlations with the QB data. Table 3 contains

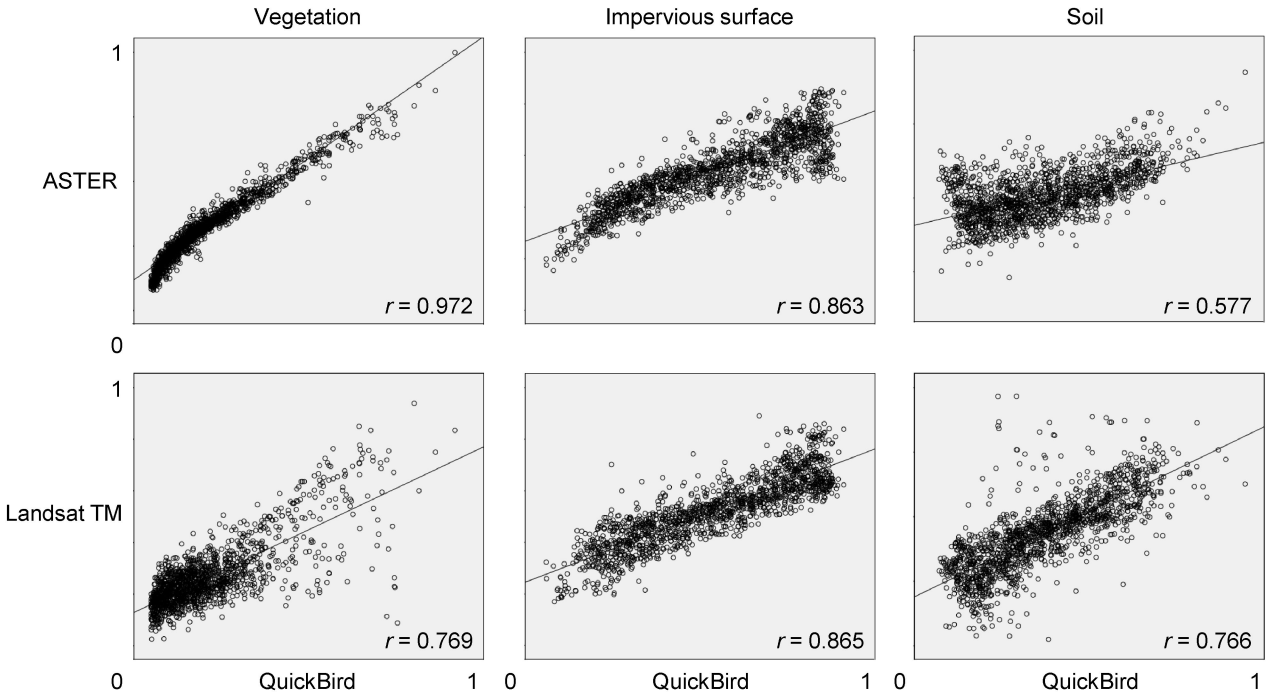

Fig. 3. Scatterplots of SMA-derived ASTER and Landsat TM land cover fractions versus proportions from the QB per-pixel classification for corresponding EAs.

Table 3. Results of Univariate, Multivariate Ordinary Least Squares, and Multivariate Spatial Error Regression Models of Vegetation and Impervious Surface Fractions on the Slum Index

	Univariate model	Multivariate OLS model		Spatial error model	
	R^2	R^2	β	Pseudo-R^2	β
ASTER SMA		0.263		0.401	
Vegetation	0.26		–0.605*		–2.707*
Impervious	0.165		–0.112*		–0.293
QuickBird per-pixel		0.254		0.396	
Vegetation	0.245		–0.372*		–1.223*
Impervious	0.202		0.155*		0.348*
Landsat TM SMA		0.186		0.388	
Vegetation	0.148		–0.289*		–2.580*
Impervious	0.119		0.217*		1.159*
ASTER per-pixel		0.155		0.373	
Vegetation	0.109		–0.167*		–0.707*
Impervious	0.137		0.269*		0.318*

* $p < 0.01$.

results of univariate analysis between the fractions and SI, as well as the initial ordinary least squares (OLS) multiple regression model.

Moran's *I* statistic is used to determine if spatial autocorrelation is present in the standardized residuals for each OLS model. The standardized residuals for all four models are positively spatially autocorrelated with *I* values corresponding to high *Z*-scores that range from 30.46 to 39.29 ($p < .01$). This reveals a spatial component to the VIS regressions that is not accounted for in the original models. Spatial heterogeneity can be expected in land cover processes, particularly with small areal units of analysis like the EA; previous work has found similarly high spatial autocorrelation in Accra's neighborhood characteristics (Weeks et al., 2007, 2010).

Results from spatial error models accounting for spatial heterogeneity at the EA scale are listed in Table 3 alongside the original univariate and multivariate results for comparison purposes; all four data sets show improved R^2 values. The Landsat TM SMA and ASTER per-pixel data yield pseudo-R^2 values that are twice as high (0.388 and 0.373, respectively) as the OLS model R^2 values. ASTER SMA and QB data again exhibit the highest R^2 (0.401 and 0.396, respectively), although with less improvement from the original model. The ASTER SMA-derived fractions are influenced less by the spatial component than the per-pixel data, which suggests the possibility that SMA may help assuage the influence of spatial autocorrelation at 15 m spatial resolution.

Results in Table 3 suggest that imagery-derived variables alone can account for up to 40% of the variation in the Slum Index at the EA level. It is also clear that the SMA-derived ASTER data are just as effective at predicting the SI values as the higher

spatial resolution QB data. The comparatively lower R^2 of the per-pixel ASTER data indicates that SMA may have some influence in creating better agreement at the differing spatial resolutions. The spatial error model also increases the agreement between image data sets relative to the original model, as differences between R^2 values are smaller when accounting for spatial autocorrelation in the residuals. Despite the differences in R^2 between data sets, the order of predicting power is the same for both the aspatial and spatial models: SMA-derived ASTER data are the strongest predictor of the SI, followed by QB, TM, and per-pixel ASTER data. Vegetation is consistently the most influential independent variable for predicting SI across all data sets, although the regression coefficients for the spatial error model are slightly lower.

Outliers are identified for both the ASTER and Landsat TM SMA standardized residual data produced by the regression in order to highlight any characteristics in the data that may create confusion in the SMA- based regressions. Of the 1,427 EAs used in the analysis, 74 EAs qualify as outliers at a 95% confidence threshold for the model based on ASTER SMA data, and 70 EAs for the TM SMA model; 60 EAs are outliers for both data sets. There are more negative outliers than positive ones, suggesting that the regression model has more problems overestimating the slum index rather than underestimating it. SI values for these EAs reveal confusion when attempting to predict extremely low values (least slum-like) and to a lesser extent high values (most slum-like). The ten lowest SI values, as well as 35 of the lowest 50, are identified as outliers in both data sets. All have extreme negative residuals, meaning that the model estimation is significantly higher than the actual SI. EAs with the eight highest SI values are also found to be outliers in both data sets.

There is no overwhelmingly distinct pattern in the VIS fractions for outlier EAs, but a few trends may be meaningful. Positively skewed outliers in the ASTER data tend to have high SI values in conjunction with high vegetation fractions relative to the mean of the entire data set, contrary to the general inverse relationship between the SI and vegetation. Some are larger EAs with lower population density along the periphery of the metropolitan area; from the QB imagery these EAs appear to be rural or industrial with little or no residential infrastructure. Non-residential areas could pose a problem for predicting slum presence using land cover such as vegetation. Peri-urban areas in particular may exhibit some of the household criteria used to calculate the slum index despite high fractions of vegetation. Patterns among negatively skewed outliers are harder to identify. A cluster of negative outliers adjacent to marshland was found in both data sets. These areas feature higher-than-average soil fractions and a complex gridded network of dirt roads visible from the QB imagery. The absence of soil in the regression equation could create confusion in these areas, but the impact is likely minimal when considering the weak correlation of soil to SI.

SUMMARY AND CONCLUSIONS

This paper explores the impact of scale on land cover fraction data, as well as the relationships between land cover and the Slum Index. High correlations indicate general agreement between VIS fractions derived from SMA of ASTER and Landsat TM imagery, and the high spatial resolution QB data. Correlations with the SI offer further evidence that the SMA-derived ASTER fractions might be an ample substitute for more expensive high spatial resolution QB imagery in an urban demographic application,

but the coarser Landsat TM data do not hold up to this standard. Vegetation fractions created from the SMA ASTER data in some instances outperform vegetation percentages from a hard QB classification in their relationship with the SI, while impervious surface fractions are similarly correlated. Yet the impact of the modifiable areal unit problem on this relationship is apparent when examining the weaker relationships of the Landsat TM fractions with the SI. Regression results reinforce the subsequent underperformance of the Landsat TM data relative to ASTER and QB. The explanatory power of VIS data appears relatively unaltered between 2.4 m and 15 m, but is adversely impacted at 30 m spatial resolution, despite the fact that SWIR bands are incorporated for TM and not the other image data sets.

It is clear from the SI correlations and regression results that the SMA classification technique used on the ASTER data has a substantial impact on the results. SMA produces fractions different from those produced by per-pixel classifiers, and these fractions are more closely related to the SI. Specifically, the SMA-derived ASTER fractions produce stronger correlations with the SI for the relevant land cover classes of vegetation and impervious surface, and are also a marginally stronger predictor of the SI. In this application, SMA is the better method for deriving proxy measures of slum conditions. These relationships remain qualified by the size of features on the ground; the variation in tree canopies, road and trail width (especially when unpaved), and other landscape features may magnify the advantages of SMA in certain parts of the city. Also, differences in the time of year and day of acquisition of different image types could account for some difference in VIS estimates, subsequent correlations between land cover fractions, and prediction of SI values. Detailed knowledge of intra-urban land cover variation is needed before adopting this approach in other urban contexts.

Finally, the correlation and regression results suggest that there is a modest relationship between vegetation and the SI. The vegetation and impervious surface fractions from both the QB and ASTER SMA are moderately correlated with the SI and are the best-performing components for predicting the SI. Vegetation is the strongest indicator of slum presence among the VIS variables and has the most potential as a predictor variable for demographic analysis. While the PPI is better at locating pure vegetation pixels (than other endmember types) and endmember selection improves by incorporating high spatial resolution imagery, it is still important to model non-vegetation land cover, as three- and four-endmember models generally yield superior vegetation fraction maps (Song, 2005). Although there is a significant link between vegetation and the SI, vegetation alone cannot confidently predict slum presence, although the use of the spatial error model does improve the results. The vegetation fraction may be more useful when used in conjunction with other covariates to create a more robust model for predicting slum presence, or when utilized as a proxy for neighborhood characteristics.

SMA is not performed on the Quickbird image in this study, and future research on the impact of sub-pixel vs. per-pixel techniques on high spatial resolution imagery could build on the results presented here. If SMA of ASTER data can produce results similar to a QB hard classification, perhaps applying SMA to the high spatial resolution imagery will yield even stronger relationships between land cover and socioeconomic variables. An object-based classification approach using the VIS framework has also shown potential for the delineation of neighborhoods of varying socioeconomic status

(Stow et al., 2007, 2010). The use of Multiple Endmember Spectral Mixture Analysis (MESMA) was considered for this analysis but was ultimately not used because the spectral libraries normally used to implement MESMA are unavailable, and creating one was beyond the scope of this work. The use of simple SMA is more suited to the parsimonious framework of this analysis, constructed with the idea that such methodologies could be used for developing cities where ground data are scarce or unreliable. SMA remains useful for this type of analysis, but it would be interesting to see how MESMA could impact the utility of land cover fraction data.

ACKNOWLEDGMENTS

This research was funded by grant number R01 HD054906 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development ("Health, Poverty and Place in Accra, Ghana," John R. Weeks, Project Director/Principal Investigator). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Child Health and Human Development or the National Institutes of Health. The authors thank two anonymous reviewers for comments that helped improve this paper.

REFERENCES

Anderson, J. R., Hardy, E. E., Roach, J. T., and R. E. Witmer, 1976, *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*, Washington, DC: U.S. Government Printing Office, U.S. Geological Survey Paper 964, 36 p.

Anselin, L., Syabri, I., and Y. Kho, 2006, "GeoDa: An Introduction to Spatial Data Analysis," *Geographical Analysis*, 38:5–22.

Baudot, Y., 2001, "A Method for the Geographical Analysis of the Population of Fast-Growing Cities in the Third World," in *Remote Sensing and Urban Analysis*, Donnay, J.-P., Barnsley, M. J., and P. A. Longley (Eds.), Boca Raton, FL: CRC Press, 249–268.

Benson, B. J. and M. D. MacKenzie, 1995, "Effects of the MAUP on Image Classification," *Geographical Systems,* 3:123–141.

Chen, K., 2002, "An Approach to Linking Remotely Sensed Data and Areal Census Data," *International Journal of Remote Sensing,* 23(1):37–48.

Cohen, J., 1982, "Set Correlation as a General Multivariate Data-Analytic Method," *Multivariate Behavioral Research*, 17:301–341.

Cross, A. M., Settle, J. J., Drake, N. A., and R. T. M. Paivinen, 1991, "Subpixel Measurement of Tropical Forest Cover Using AVHRR Data," *International Journal of Remote Sensing,* 12(5):1119–1129.

Cushnie, J. L., 1987, "The Interactive Effect of Spatial Resolution and Degree of Internal Variability within Land-Cover Types on Classification Accuracies," *International Journal of Remote Sensing*, 8(1):15–29.

Fisher, P., 1997, "The Pixel: A Snare and a Delusion," *International Journal of Remote Sensing*, 18(3):679–685.

Hair, J. F., Jr., Anderson, R. E., Tatham, R. L., and W. C. Black, 1998, *Multivariate Data Analysis*, Englewood Cliffs, NJ: Prentice Hall, 768 p.

Herold, M., Gardner, M., Hadely, B., and D. A. Roberts, 2002, "The Spectral Dimension in Urban Land Cover Mapping from High-Resolution Optical Remote Sensing Data," in *3rd International Symposium Remote Sensing of Urban Areas*, 11–12 June 2002, Istanbul Technical University, Istanbul, Turkey, 77–91.

Iverson, L. R., Cook, E. A., and R. L. Graham, 1989, "A Technique for Extrapolating and Validating Forest Cover Across Large Regions," *International Journal of Remote Sensing,* 10(11):1805–1812.

Johansen, K., Phinn, S., and C. Witte, 2010, "Mapping of Riparian Zone Attributes Using Discrete Return LiDAR, QuickBird, and SPOT-5 Imagery: Assessing Accuracy and Costs," *Remote Sensing of Environment,* 114:2679–2691.

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(9):1053–1062.

Marceau, D. J., 1999, "The Scale Issue in Social and Natural Sciences," *Canadian Journal of Remote Sensing*, 25(4):347–356.

Marceau, D. J., Gratton, D. J., Fournier, R., and J. P. Fortin, 1994, "Remote Sensing and the Measurement of Geographical Entities in a Forested Environment; Part 1: The Scale and Aggregation Problem," *Remote Sensing of Environment,* 49:93–104.

Marceau, D. J. and G. J. Hay, 1999, "Remote Sensing Contributions to the Scale Issue," *Canadian Journal of Remote Sensing,* 25(4):357–366.

National Research Council, 2007, *Tools and Methods for Estimating Population at Risk from Natural Disasters and Complex Humanitarian Crises*, Washington, DC: National Academies Press.

Openshaw, S., 1984, *The Modifiable Areal Unit Problem*, Norwich, UK: Geobooks, 41 p.

Pellow, D., 2002, *Landlords and Lodgers: Spatial Organization in an Accra Community*, Westport, CT: Praeger Publishers, 280 p.

Phinn, S., Stanford, M., Scarth, P., Murray, A. T., and P. T. Shyy, 2002, "Monitoring the Composition of Urban Environments Based on the Vegetation-Impervious Surface-Soil (VIS) Model by Subpixel Analysis Techniques," *International Journal of Remote Sensing*, 23(20):4131–4153.

Rashed, T., Weeks, J. R., Gadalla, M. S., and A. G. Hill, 2001, "Revealing the Anatomy of Cities through Spectral Mixture Analysis of Multispectral Satellite Imagery: A Case Study of the Greater Cairo Region, Egypt," *Geocarto International*, 16(4):5–15.

Rashed, T., Weeks, J. R., Stow, D., and D. Fugate, 2005, "Measuring Temporal Compositions of Urban Morphology Through Spectral Mixture Analysis: Toward a Soft Approach to Change Analysis in Crowded Cities," *International Journal of Remote Sensing*, 26(4):699–718.

Ridd, M., 1995, "Exploring a V-I-S (Vegetation-Impervious Surface-Soil) Model for Urban Ecosystem Analysis Through Remote Sensing: Comparative Anatomy of Cities," *International Journal of Remote Sensing*, 16(12):2165–2185.

Rindfuss, R. R. and P. C. Stern, 1998, "Linking Remote Sensing and Social Science: The Needs and the Challenge," in *People and Pixels*, Liverman, D., Moran, E. F., Rindfuss, R. R., and P. C. Stern (Eds.), Washington, DC: National Academy Press.

Rogerson, P. A., 2001, *Statistical Methods for Geography.* London, UK: SAGE Publications Ltd., 320 p.

Small, C. and J. W. T. Lu, 2006, "Estimation and Vicarious Validation of Urban Vegetation Abundance by Spectral Mixture Analysis," *Remote Sensing of Environment*, 100:441–456.

Song, C., 2005, "Spectral Mixture Analysis for Subpixel Vegetation Fractions in the Urban Environment: How to Incorporate Endmember Variability?," *Remote Sensing of Environment*, 95:248–263.

Stoler, J., Weeks, J. R., Getis, A., and A. G. Hill, 2009, "Distance Threshold for the Effect of Urban Agriculture on Elevated Self-Reported Malaria Prevalence in Accra, Ghana," *American Journal of Tropical Medicine and Hygiene*, 80(4):547–554.

Stow, D. A., Lippitt, C. D., and J. R. Weeks, 2010, "Geographic Object-based Delineation of Neighborhoods of Accra, Ghana using QuickBird Satellite Imagery," *Photogrammetric Engineering and Remote Sensing,* 76(8):907–914.

Stow, D. A., Lopez, A., Lippitt, C., Hinton, S., and J. R. Weeks, 2007, "Object-Based Classification of Residential Land Use within Accra, Ghana Based on QuickBird Satellite Data," *International Journal of Remote Sensing*, 28(22):5167–5173.

United Nations, 2010, *World Urbanization Prospects: The 2009 Revision*, New York, NY: United Nations Population Division, Department of Economic and Social Affairs.

Van Der Meer, F. and S. M. De Jong, 2000, "Improving the Results of Spectral Unmixing of Thematic Mapper Imagery by Enhancing the Orthogonality of Endmembers," *International Journal of Remote Sensing*, 21(15):2781–2797.

Ward, D., Phinn, S. R., and A. T. Murray, 2000, "Monitoring Growth in Rapidly Urbanizing Areas Using Remotely Sensed Data," *The Professional Geographer*, 52(3):371–385.

Weeks, J. R., Getis, A., Hill, A. G., Agyei-Mensah, S., and D. Rain, 2010, "Neighborhoods and Fertility in Accra, Ghana: An AMOEBA-Based Approach," *Annals of the Association of American Geographers*, 100(3):558–578.

Weeks, J. R., Hill, A. G., Getis, A., and D. A. Stow, 2006, "Ethnic Residential Patterns as Predictors of Intra-urban Child Mortality Inequality in Accra, Ghana," *Urban Geography,* 27(6):526–548.

Weeks, J. R., Hill, A. G., Stow, D. A., Getis, A., and D. Fugate, 2007, "Can We Spot a Neighborhood From the Ground? Defining Neighborhood Structure in Accra, Ghana," *Geojournal,* 69:9–22.

Wu, C. and A. T. Murray, 2003, "Estimating Impervious Surface Distribution by Spectral Mixture Analysis," *Remote Sensing of Environment,* 84:493–505.

Wu, C. and A. T. Murray, 2005, "A Cokriging Method for Estimating Population Density in Urban Areas," *Computers, Environment and Urban Systems,* 29:558–579.

Zhang, J. and G. M. Foody, 1998, "A Fuzzy Classification of Sub-urban Land Cover from Remotely Sensed Imagery," *International Journal of Remote Sensing*, 19(14):2721–2738.