Working Paper No. 178

Viewing society from space: Image-based sociocultural prediction models

by John M. Irvine, Richard J. Wood, and Payden McBee
| December 2017

Working Paper No. 178

Viewing society from space: Image-based sociocultural prediction models

by John M. Irvine, Richard J. Wood, and Payden McBee | December 2017

John M. Irvine (corresponding author) is chief scientist for data analytics at Charles Stark Draper Laboratory. Email: jirvine@draper.com.
Richard J. Wood is a principal member of the technical staff at Charles Stark Draper Laboratory.
Payden McBee is a fellow at Charles Stark Draper Laboratory.

Abstract

Applying new analytic methods to imagery data offers the potential to dramatically expand the information available for human geography. Satellite imagery can yield detailed local information about physical infrastructure, which we exploit for analysis of local socioeconomic conditions. Combining automated processing of satellite imagery with advanced modeling techniques provides a method for inferring measures of well-being, governance, and related sociocultural attributes from satellite imagery. This research represents a new approach to human geography by explicitly analyzing the relationship between observable physical attributes and societal characteristics and institutions of the region. Through analysis of commercial satellite imagery and Afrobarometer survey data, we have developed and demonstrated models for selected countries in sub-Saharan Africa (Botswana, Kenya, Zimbabwe). The findings show the potential for predicting people's attitudes about the economy, security, leadership, social involvement, and related questions, based only on imagery-derived information. The approach pursued here builds on earlier work in Afghanistan. Models for predicting economic attributes (presence of key infrastructure, attitudes about the economy, perceptions of crime, and outlook toward the future) all exhibit statistically significant performance. Although these results are encouraging, several avenues for advancement and improvement are proposed. Initial analysis of new methods for image processing and feature extraction have identified several avenues for promising enhancements.

Acknowledgements

This research was supported in part by the National Geospatial-Intelligence Agency (NGA) under the GEOINT Analysis Program and in part by Charles Stark Draper Laboratory (Draper) under the Draper Fellows Program. All survey data were provided by Afrobarometer. The commercial satellite imagery from DigitalGlobe was provided by NGA. The opinions expressed here are those of the authors and do not necessarily reflect the views of Draper, NGA, DigitalGlobe, or Afrobarometer.

I. Introduction

Globalization and growth in international commerce have elevated the importance of understanding diverse cultures and societies. Applying new analytic methods to imagery data offers the potential to dramatically expand the information available for human geography. Satellite imagery can acquire detailed local information about physical infrastructure, which is the basis for estimating local socioeconomic conditions. A recent study in Afghanistan indicates tremendous potential for exploiting overhead imagery to characterize economic well-being, governance, social capital, and related sociocultural factors (Irvine & Lepanto, 2014). To realize this potential for imagery-based analysis to provide indicators of these phenomena, we conducted the study presented here. The goal is to develop and demonstrate methods to predict or estimate economic, political, and sociocultural characteristics at a fine level of geospatial detail through analysis of commercial satellite imagery. This study builds on recent work performed with similar data for Afghanistan.

Previous work focused on rural villages in Afghanistan and exhibited excellent results. Features extracted from commercial satellite imagery were predictive of local survey responses to numerous questions about the economy, attitudes toward governance and security, engagement in community projects, and related sociocultural factors. The broad-based applicability of these methods, however, required further investigation. The focus of this effort is to explore similar analysis in other parts of the world, namely in Africa (Botswana, Kenya, and Zimbabwe), to assess the extensibility and generalizability of these methods. The methods developed here can enable the automated or semi-automated production of geospatially localized estimates of key economic, political, and social indicators, addressing a critical need for human geography.

II. Background

The application of remote sensing and geographic information systems (GIS) to the social sciences is an emerging research area (Blumberg & Jacobson, 1997; Goodchild, Anselin, Appelbaum, & Harthorn, 2000; Hall, 2010; Taubenböck et al., 2009; Crews, 2009; National Research Council, 1998). Recognizing that people's behaviour and values shape, and are shaped by, the environment in which they live, researchers have explored a number of issues, including sociocultural and economic attributes, ethnography, and land use (Fox, Rindfuss, Walsh, & Mishra, 2003; Jiang, 2003). Through prior work we developed and demonstrated semi-automated algorithms for feature extraction from images that fed into modeling and estimation of sociocultural phenomena. These methods extract both spatial and spectral features from the imagery, using supervised learning, to yield a rich set of image descriptors. The resulting models for rural Afghanistan successfully predicted a number of indicators of local economic and social conditions (Irvine, Lepanto, Regan, & Young, 2012; Irvine et al., 2013; Irvine & Lepanto, 2014).

Physical structures provide imagery observables that indicate local economic and social conditions. Some basic principles that have emerged from empirical studies in the social sciences suggest relations between the economic indicators and social/political factors:

- Economic inequality negatively affects the establishment and maintenance of democratic practice (Boix, 2003; Acemoglu & Robinson, 2009).
- Ethnic diversity reduces public goods provision (Habyarimana, Humphreys, Posner, & Weinstein, 2007).
- Lower gross domestic product (GDP) negatively affects democracy and public goods provision (Kennedy, 2010), whereas democratic practice increases overall wealth (Rodrik, 2000).
- Lower concentration of households leads to lower production of public goods (Rauk, 1993).

Copyright ©Afrobarometer 2017

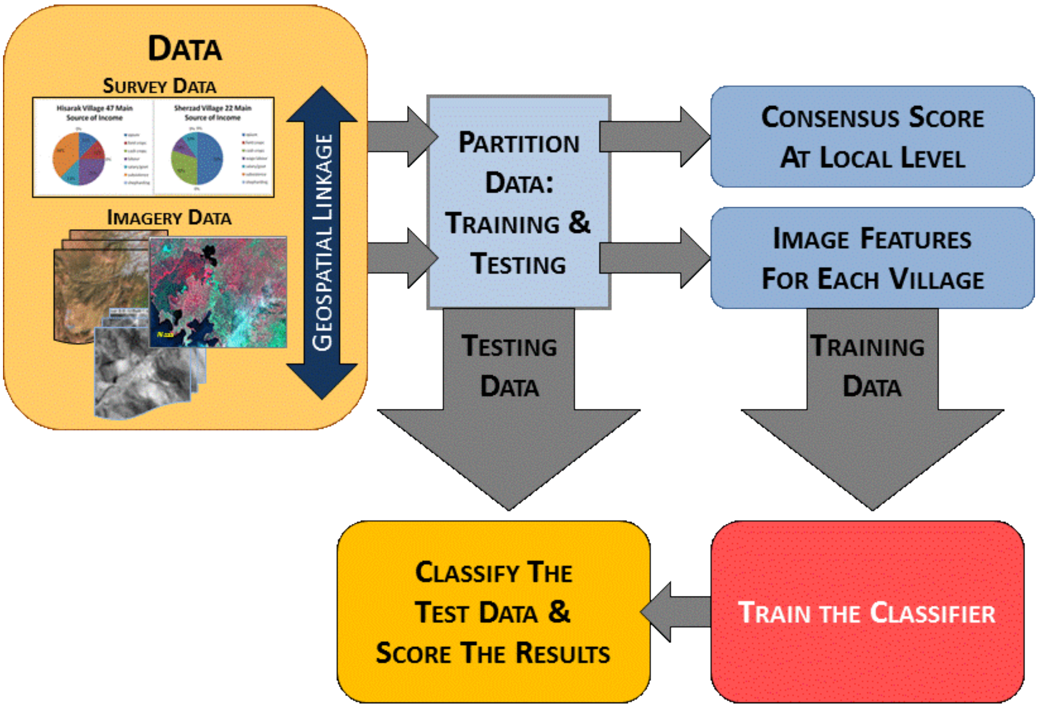
- High social capital has been linked to good governance (Bowles & Gintis, 2002).

The functioning of society depends on the interplay among social organizations, institutions, and physical structures. The physical structures constitute the features observable with imagery. Our research in Afghanistan yielded an initial methodology for predicting economic, political, social, and cultural indicators from commercial satellite imagery. The results demonstrated that predictive classification accuracy for binary responses across 76 survey-based indicators ranged from 70% to 95% on the test data. The models showed that imagery and geospatial data can predict relevant social indicators much better than chance.

III. Technical approach

The approach combines an understanding of political, sociocultural, and economic theory with rich sources of survey data and overhead imagery to analyze the relationships between measurements acquired from direct surveys and features inferred from remotely-sensed imagery. The goal is to develop models that predict the values of specific indicators using the features extracted from the satellite imagery. The full approach consists of the steps illustrated in Figure 1.

Figure 1: Overview of the model development process



1. Analysis of survey data: Analysis of the survey questions and responses produces indicators of local attitudes about economic well-being, attitudes toward democratic ideals and practice, indicators of social capital, and concerns about crime and security.
2. Imagery analysis: Using image processing techniques to extract relevant features from the imagery, we construct measures of local conditions.
3. Model development: Statistical analysis of the relationships between the survey-based indicators and imagery-derived features yields the initial models. The overhead imagery covers the same geographic region as the survey data, and the

Copyright ©Afrobarometer 2017

geolocation of survey responses is the link between the two data sources. Linking the two data sources depends on the fidelity of the geospatial information associated with the two data types. For the imagery, geolocation is straightforward: The corner coordinates (latitude and longitude) for each image appear in the sensor metadata, and using standard photogrammetric techniques produces a precise location for each region in the image. Determining the location data for the surveys was more challenging. The survey data included the village or city and neighbourhood names. Using the gazetteer associated with Google Earth, we looked up each name and extracted the center coordinates provided by Google Earth. This process introduced some ambiguity in major urban areas, as discussed below.

4. Model validation and analysis: Model validation requires a separate set of survey and imagery data that is held back (i.e. sequestered) during the model development process. Using this new set of data, we computed survey-based indicators and imagery-derived features. Comparing the observed survey-based indicators to model predictions provides an empirical measure of performance.

IV. Data sources

Implementing the methodology requires two sources of data that are linked through common geospatial tagging. Commercial satellite imagery provides the source we are exploiting to characterize socioeconomic conditions. Survey data collected over the same regions during a comparable time frame constitutes the "ground truth" that we seek to predict from the imagery. The geospatial locations of the images and survey responses form the connection to associate the two data sources.

IV.A Survey data

A basic approach to understanding individuals and societies consists of asking them questions directly. A conversation with one individual is termed an interview. While interviews can be revealing, the information is specific to the person being interviewed. Surveys have many advantages. They offer a direct measurement of attitudes and opinions, the methodology has a high degree of transparency and repeatability, the data acquired are amenable to analysis using standard statistical tools, and the results and findings can be represented quantitatively. Some of the drawbacks arise from the costs of conducting surveys, the challenges of designing good survey instruments, and the difficulty of performing surveys in some parts of the world. Not all respondents may feel comfortable giving honest responses to certain types of questions, either because of cultural norms or because of safety concerns. In some cases, researchers can devise questionnaires that address sensitive issues in ways that make the respondent more comfortable; in other cases, it may be impossible to obtain responses to certain questions. For example, in a country with a regime known for suppressing dissent, survey respondents are unlikely to give honest answers to questions about support for the government. An alternative to using survey data is to rely other sources. For example, the Living Standards Measurement Study produced by the World Bank (2017) offers a rich source of information. But it is not available for the countries studied here. Other sources of data, including government-produced statistics, are generally not available at a local geographic level and can be years out of date (Bagus, 2006; Jerven, 2014).

For this study, we used an extensive set of public opinion data collected in sub-Saharan Africa by the Afrobarometer network in its Round 4 (2008/2009) surveys. Afrobarometer is a pan-African research collective that conducts surveys in more than 30 countries. Core members include the Ghana Center for Democratic Development, the Institute for Empirical Research in Political Economy (Benin), the University of Nairobi's Institute for Development Studies, and the Institute for Justice and Reconciliation (South Africa), with support from Michigan State University and the University of Cape Town Institute for Democracy, Citizenship and Public Policy in Africa. Each Afrobarometer survey collects data about individual attitudes and behaviour, including indicators relevant to developing societies.

Copyright ©Afrobarometer 2017

Additional data include the presence/absence of infrastructure. Core topics addressed in each survey round are summarized in Table 1. The survey data provide a rich portrait of societal attitudes across countries, multiple regions within each country, socio-demographic groups, and points in time.

Table 1: Core topics addressed in Afrobarometer surveys

Democracy	Popular understanding of, support for, and satisfaction with democracy, as well as any desire to return to, or experiment with, authoritarian alternatives.
Elections	Participation in campaigns and elections. Citizens' voting intentions and their opinions on the quality of electoral processes.
Governance	The demand for, and satisfaction with, effective, accountable, and clean government; judgments of overall governance performance and social service delivery.
Public services	The availability of public services and how often it is accessed. Public services include piped water, health clinic, cell phone service and postal systems.
Macro-economics and markets	Assessments of national and personal economic and living conditions. Evaluations of government performance in economy management and creating jobs.
Poverty	How often do individuals experience shortages of basic essentials – food, water, medical care – in their daily lives? Indicators of basic living conditions.
Social capital	Whom do people trust? To what extent do they rely on informal networks and associations? What are their evaluations of the trustworthiness of various institutions?
Gender equality	Women's position in society. Should women have the same rights as men? Should there be more female leaders in politics and public institutions?
Conflict and crime	How safe do people feel? What has been their experience with crime and violence? Do they report crimes to the police?
Participation	The extent to which ordinary people join in development efforts, comply with the laws of the land, vote in elections, contact elected representatives, and engage in protest. The quality of electoral representation.
Identity	How do people see themselves in relation to ethnic and class identities? Does a shared sense of national identity exist?
Tolerance	How accepting are people of those who are socially or politically different from themselves?

IV.B Imagery data

Several recent studies have demonstrated that imagery collected from space-based or airborne imaging assets can reveal useful information about societies and the environment in which they live. The LandScan Project at Oak Ridge National Laboratory has developed methods for estimating populations and population distribution on a global basis (Landscan; Bhaduri, Bright, Coleman, & Urban, 2007; Cheriyadat, Bright, Bhaduri, & Potere, 2007; Vijayaraj, Bright, & Bhaduri, 2007; Vijayaraj et al., 2008). This program has demonstrated the

Copyright ©Afrobarometer 2017

practical application of remote sensing methods for mapping human settlements and analyzing population movements over time. A recent study of the urban landscape in Guatemala (Owen, 2012) represents a more focused application of remote sensing. Owen's work made an important contribution through an assessment that distinguished informal (slum) and formal (planned) settlements using high-resolution imagery. Min, Agnew, Gillespie, and Gonzalez (2008) made a novel use of "nightlights"[1] data to assess political and military activity remotely. More recently, several researchers have explored methods for assessing poverty from overhead data (Henderson, Storeygard, & Weil, 2012; Jean et al., 2016).

Social science research suggests several ways in which local attitudes and behaviours may correspond to phenomena observable in overhead imagery. Although this set is not exhaustive, certain issues have a foundation in social science research and a logical connection to observable phenomena (Table 2). The theory suggests natural hypotheses to examine in three areas: income and economic development, centrality and decision authority, and social capital. First, overhead imagery reveals numerous indicators of economic status (housing, vehicles, crop land, livestock, and infrastructure). Studies have explored the relationship between remote sensing data and the economy (Elvidge et al., 1997; Irvine et al., 2013). Second, both higher income and equitable distribution of income are associated with good governance. Observables associated with economic well-being, including measures of wealth distribution, also serve as relevant indicators of governance. Transportation and communication infrastructure provides indicators of expected levels of social interactions. Third, high social capital has been linked to good governance (Bowles & Gintis, 2002; Bardhan, Bowles, & Gintis, 2000), but observing meaningful measures that correlate with social capital poses a challenge. Evidence of economic growth can be associated with higher levels of social capital (Knack & Keefer, 1997). Although social cohesion and connectedness cannot be measured directly, durable institutions (e.g. schools, places of worship) and infrastructure (e.g. roads, cell towers) are indirect indicators of social connectedness.

Table 2: Research issues and potential observables to explore

Research issue	Potential observables
Income and economic development:	House sizes: average size, range of sizes Presence of motor vehicles Physical infrastructure Agriculture: extent and mix of cultivation, crop health, presence and extent of livestock
Centrality and decision authority	Road network, lines of communication Physical infrastructure (bridges, paved roads, schools, mosques)
Social capital	Community infrastructure and prevalence of meeting places and institutions (schools, places of worship) Communications infrastructure, roads

Cultural, social, and economic factors critical to understanding societal attitudes are associated with specific phenomena observable from overhead imagery. Distinguishing among industrial, commercial, and residential areas, for example, is a standard use for imagery (Harvey et al., 2002; Harvey & Theiler, 2004; O'Brien & Irvine, 2004; Jenson & Cowan, 1999). Researchers can also extract measures of socioeconomic status (e.g. house size, crop

[1] The Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS) captures night imagery of the Earth, which can reveal lighting patterns arising from human settlements.

Copyright ©Afrobarometer 2017

area, and crop vigor, presence of vehicles) from high-quality imagery. Other factors can be inferred from indicators derived from the imagery, such as population density, access to improved roads, distances to commercial and governmental centers, and attributes of communities. Patterns that emerge from the correlation of the geospatial analysis with survey data can suggest phenomena observable in imagery that serve as reasonable surrogates for the direct measurements of public opinion (Irvine, et al., 2012; Irvine, et al., 2013). Researchers can apply various image-processing techniques to extract and identify specific indicators from the imagery (Table 3).

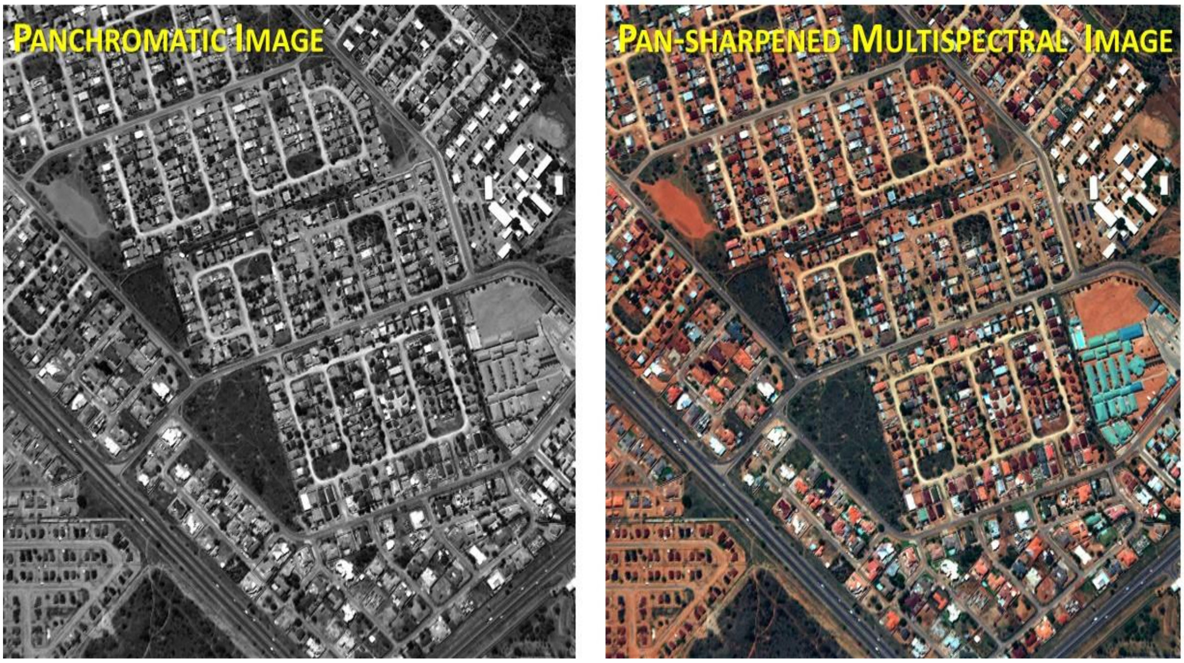
Table 3: Illustrative imagery observables and derived indicators

Feature class	Observables	Derived features
Land cover	Arable land Area under cultivation Crop health	Level of commercial agriculture Expected food supply
Buildings	Sizes of buildings Number of types of buildings Size of residential buildings Presence of a "large" building	Degree of urbanization or industrialization Population density and distribution Average size of residential buildings Indication of a mosque, school, or community building
Lines of communication	Road network (paved and unimproved) Other transportation Numbers and types of vehicles Distance from major commercial centers	Level of communication Infrastructure to support local commerce Access to transportation

Panchromatic imagery, having finer spatial resolution, can reveal details about the shapes of objects. Multispectral imagery has coarser resolution but includes data in four spectral regions (red, green, blue, and near-infrared). This spectral information can reveal material properties, such as the types of roads, buildings, crops, and vegetation. Panchromatic sharpening (Alparone, Wald, Chanussot, Bruce, & Thomas, 2007; Lee & Lee, 2010) can combine the two images to produce a fine-resolution multispectral image (Figure 2).

An important consideration is the level of detail observable in an image. Table 4 presents the important sensor parameters for the commercial systems. These satellites are generally capable of acquiring panchromatic imagery with spatial resolution (ground sample distance) of better than 1 meter. They can also collect four-band multispectral imagery that provides some information for material identification and analysis of vegetation. Pan sharpening the multispectral product yields a high-quality image for detailed exploitation. The intelligence community has promulgated a standard known as the National Imagery Interpretability Ratings Scale (NIIRS), which quantifies the interpretability of imagery (Leachtenauer & Driggers, 2001; Irvine, 2003). Both recent user studies and the general image quality equation for predicting NIIRS from sensor parameters indicate that current commercial systems collect imagery that is NIIRS 5 or better (Leachtenauer, Malila, Irvine, Colburn, & Salvaggio, 1997). For example, Figure 3 presents an image collected by GeoEye-1 over Pakistan. Note the level of detail evident in the scene. One can clearly see vehicles on the roads; identify buildings and associated compounds; delineate the extent of crops, forests, and other land cover; and analyze building size and configuration. In short, the commercial sensors can support the types of analysis proposed in this project.

Figure 2: Illustrative example of panchromatic image and pan-sharpened multispectral image



Imagery from DigitalGlobe provided by National Geospatial-Intelligence Agency (NGA)

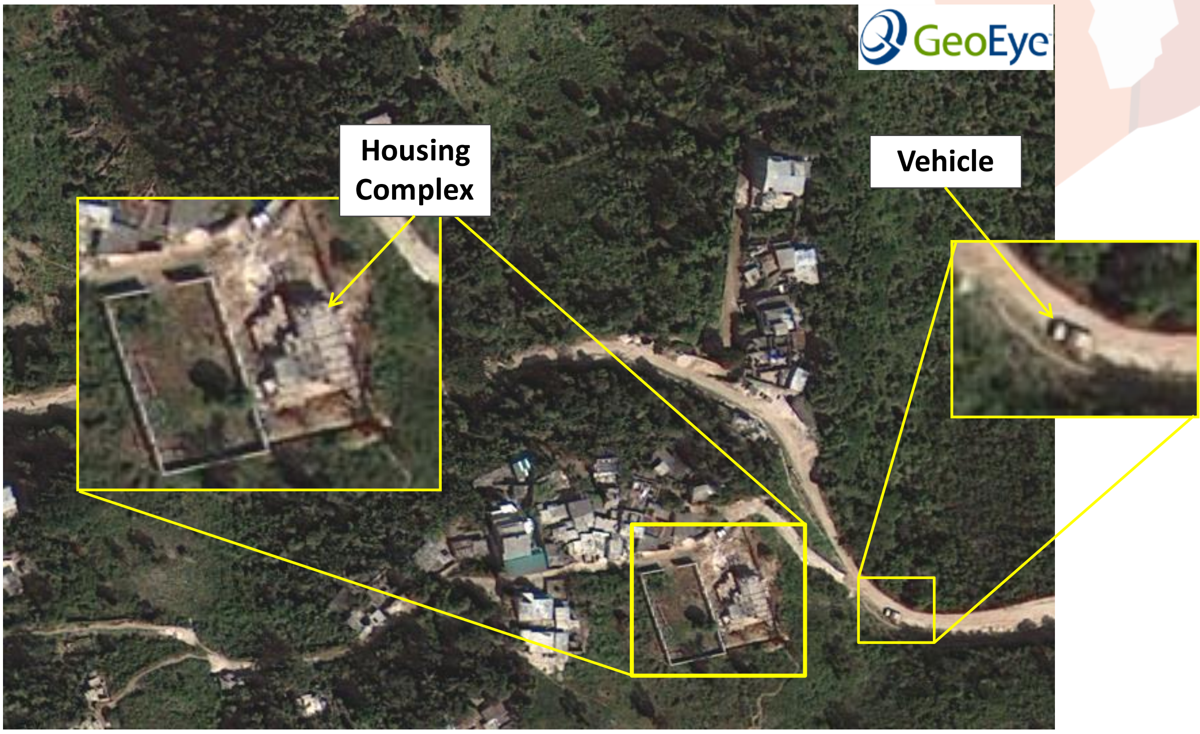
Table 4: Sensor parameters for commercial imaging satellites

Parameters	GeoEye-1	IKONOS	Quickbird	WorldView-1
Ground sample distance (GSD) at nadir for panchromatic sensor (meters)	0.41 x 0.41	0.82 x 3.2	0.65 x 0.65	0.5 x 0.5
GSD at nadir for multispectral sensor (meters)	1.65 x 1.65	3.3 x 3.3	2.62	
Spectral range (nm)	450-800	526-929	430-918	400-900
Blue	450-510	445-516	430-545	
Green	510-580	506-595	466-620	
Red	655-690	632-698	590-710	
Near infrared (IR)	780-920	757-853	715-918	
Swath width at nadir (km)	15.2	11.3	18	17.7

The National Geospatial-Intelligence Agency (NGA) provided all of the commercial satellite imagery for this project, including approximately 40 terabytes of imagery data for the study areas in sub-Saharan Africa. Both panchromatic and multispectral data were provided; the multispectral data were often provided as separate files for each band, requiring a significant effort to reassemble the data into four-band multispectral images.

Copyright ©Afrobarometer 2017

Figure 3: GeoEye-1 imagery from Pakistan



Imagery from DigitalGlobe provided by NGA

V. Analysis

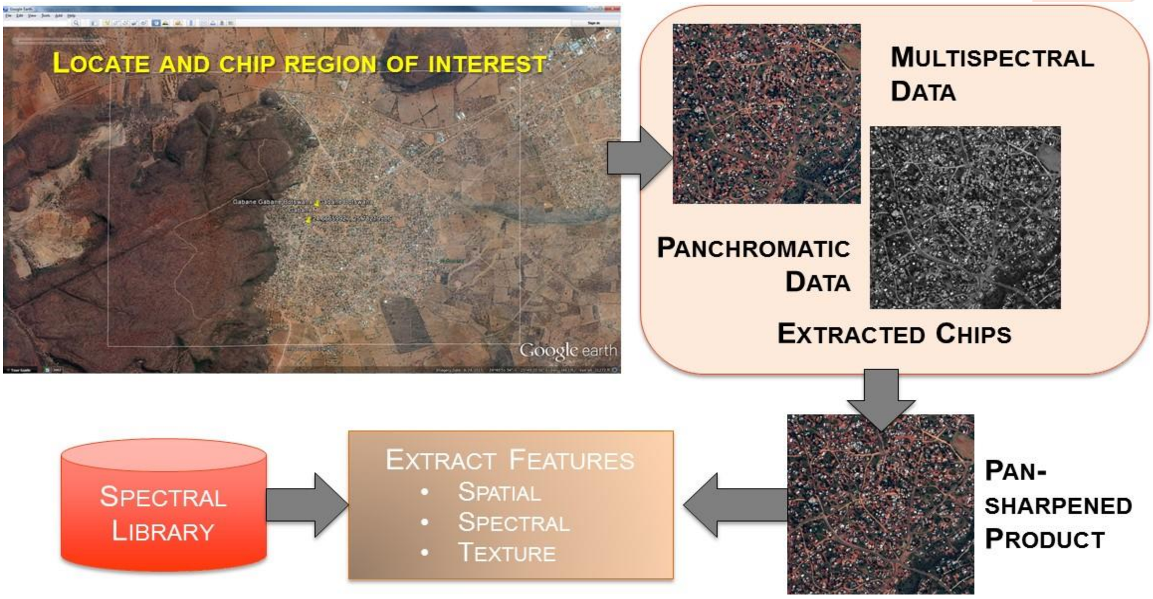
A range of image-processing methods are available for extracting traditional geospatial features (roads, buildings, land cover, water bodies) from overhead imagery (O'Brien & Irvine, 2004). Much of this research focuses on land use, land cover, and similar environmental issues for which remote sensing is ideal.

V.A Image processing and feature extraction

The first step in processing the commercial satellite imagery was to identify and select the sub-regions that coincided with the survey data. The multistep process generated pan-sharpened multispectral chips over the geographic areas corresponding to the Afrobarometer surveys (Figure 4). The process began with geolocating the villages or neighbourhoods identified in the survey data. Using the latitude and longitude for these locations, one-kilometer-square image chips were extracted from both the panchromatic images and the multispectral images. These image chips were then co-registered and processed jointly to produce the pan-sharpened products. A mix of spatial and spectral analysis extracted features related to the size and shape of the cultural features and built-up areas, the health of crops and other vegetation, material properties of the buildings and roads, access to roads and water, and other physical observables as described below.

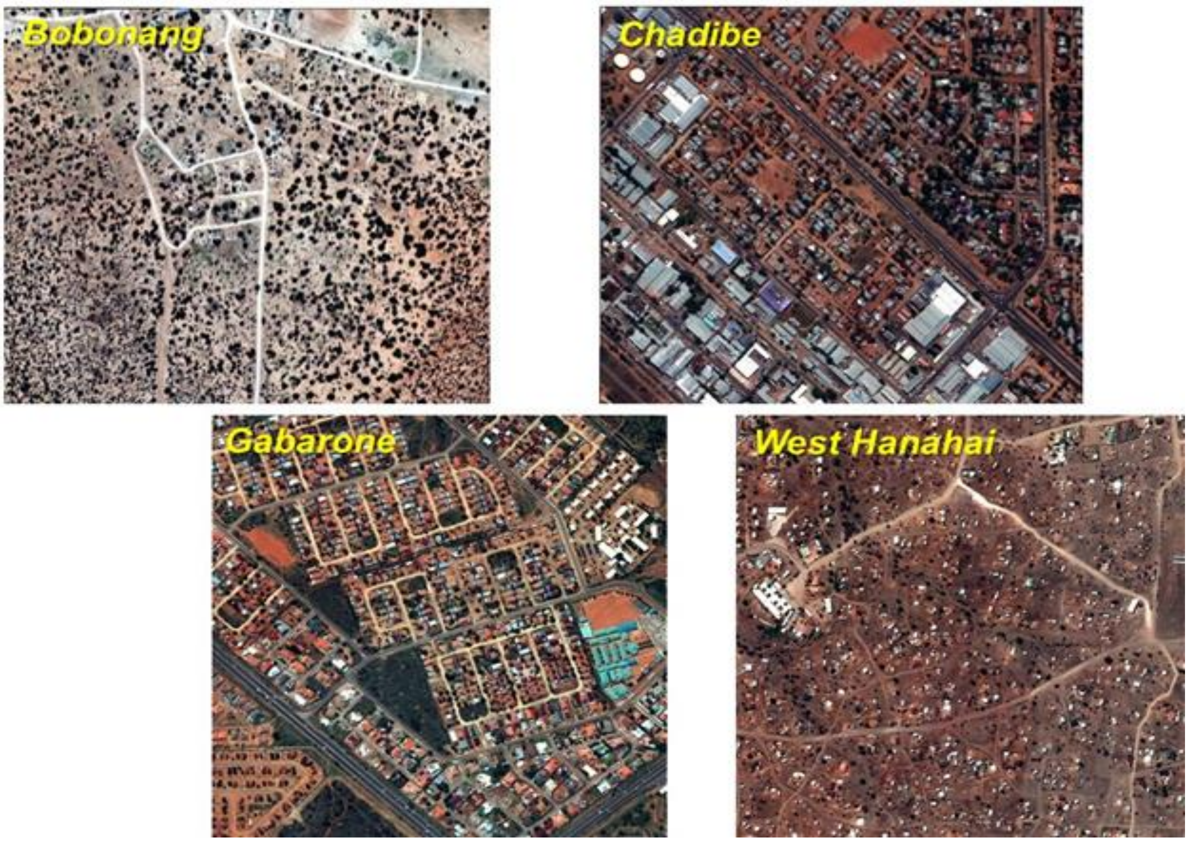
The image chipping, registration, and pan-sharpening produced a set of one-kilometer-square images centered over the locations identified in the survey data. The wide range of conditions across these locations is evident from the imagery (Figure 5). A variety of processing methods extracted features from the imagery to characterize the spatial and spectral attributes of the data. Spectral features included both relative measures and features based on spectral libraries representing known materials. The vegetation analysis is an example of a relative spectral feature. To identify regions of vegetation and assess their health, we computed the normalized difference vegetation index (NDVI), which indicates the presence and health of vegetation (Kriegler, Malila, Nalepka, & Richardson, 1969). This index of plant "greenness" or photosynthetic activity is a standard measure used by the remote sensing community (Figure 6). The mean value over an image chip indicates local vegetation extent and health, which becomes a feature used in our models.

Figure 4: Image-processing overview



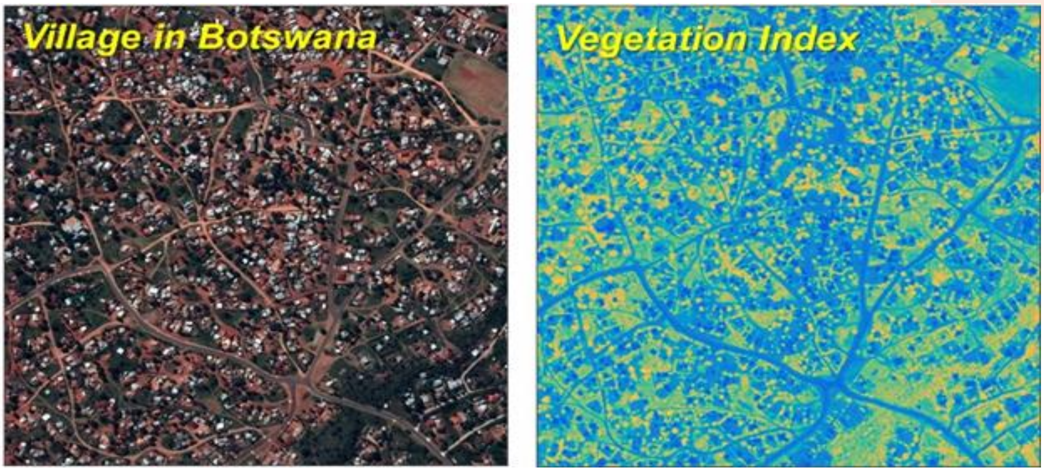
Imagery from DigitalGlobe provided by NGA

Figure 5: Four pan-sharpened chips for different locations in Botswana showing the range of urbanization, population density, vegetation, and spatial organization



Imagery from DigitalGlobe provided by NGA

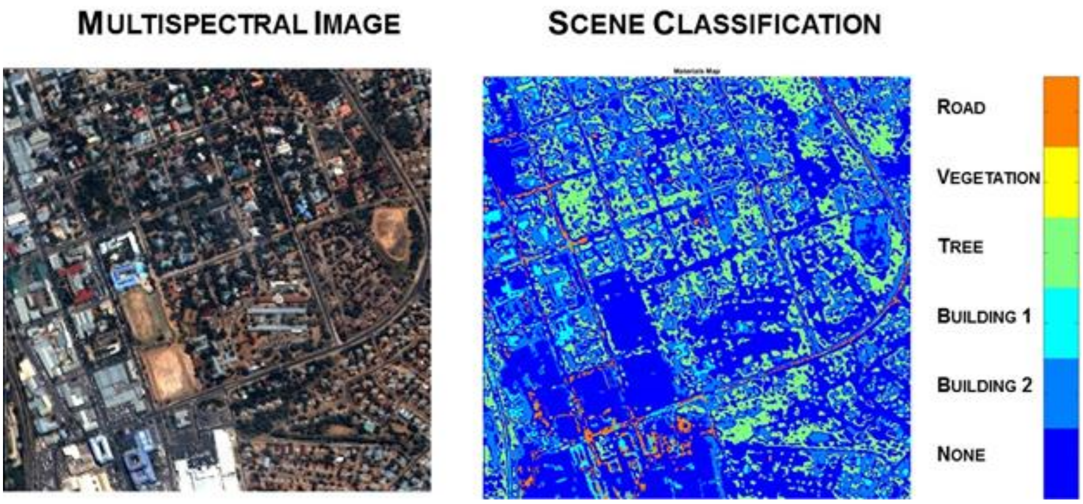
Figure 6: Pan-sharpened chip and NDVI computed for the same region



Imagery from DigitalGlobe provided by NGA

The second type of spectral analysis compares the imagery data to known spectral signatures to classify the materials in the scene. We developed a small spectral library by selecting samples from our image set and computed the average spectra from these samples. Materials of interest included roads and roofing materials. To identify the materials within a new image, we computed the spectral angle between the observed pixel and the spectra in our library. This spectral angle mapping has been widely applied to classification problems for remote sensing data (Kruse et al., 1993; Rashmi, Addamani, Venkat, & Ravikiran, 2014). Using our limited spectral library, the technique yields approximate classifications (Figure 7), but this may be sufficient to provide a statistical characterization of a region.

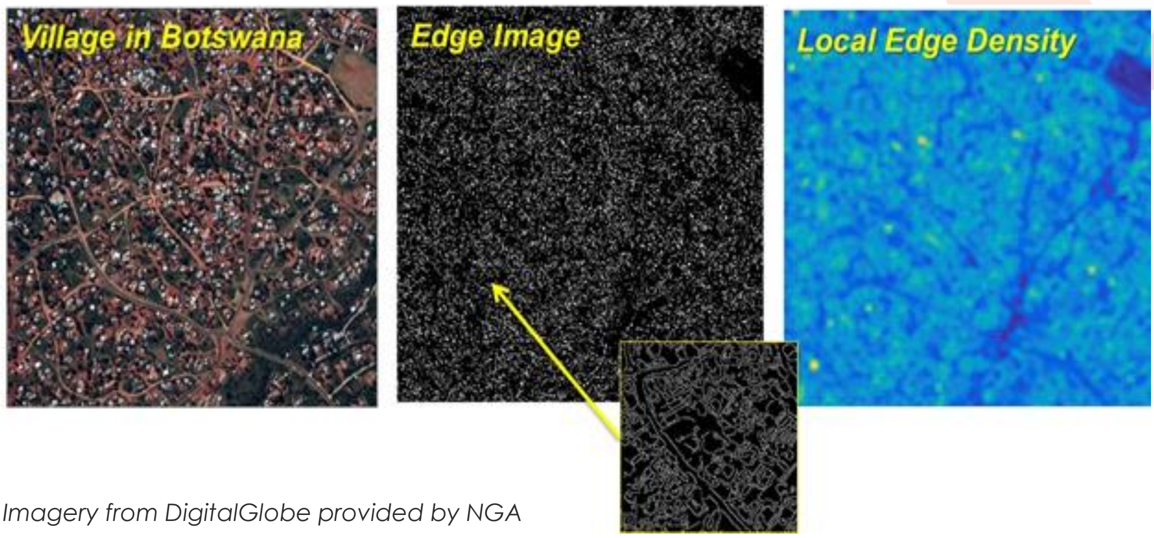
Figure 7: Scene classification using spectral angle mapping with five classes



Imagery from DigitalGlobe provided by NGA

Turning to spatial processing, we extracted a variety of features based on edge density and the spatial covariance structure of the image. Generally, human-made structures produce more sharp features in an image than natural vegetation and terrain features. Regions with a large number of edges are more likely to be a concentration of buildings. Edge density is, therefore, an indicator of building density (Figure 8). Features such as average and standard deviation of the edge density for an image help to characterize the nature of the human-made structures in a region.

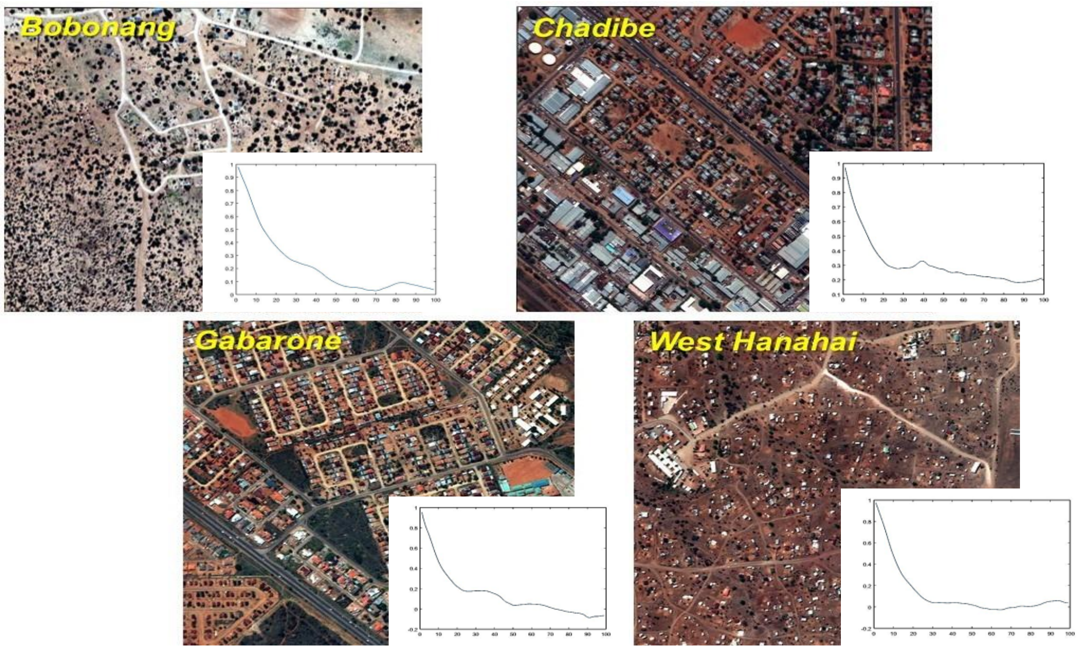
Figure 8: Density and variability of edges indicating the concentration and distribution of human-made structures



Imagery from DigitalGlobe provided by NGA

A second class of spatial features indicates the regularity of building patterns. The spatial covariance across an image suggests cyclic behaviour in urban areas that follow a grid pattern. Chaotic patterns, such as undocumented settlements in large areas, lack this systematic spatial pattern, and the covariance function decays rapidly in a manner similar to rural areas. The regular or cyclic behaviour is more common with middle-class neighbourhoods, whereas very wealthy areas show longer and less periodic behaviour (Figure 9). To extract features that characterize the spatial structures, we estimated both the horizontal and vertical covariance functions. Polynomial models of degree 3 are fit to these functions, and the coefficients of these models, which indicate the nature of the spatial variation, become the features used in subsequent analysis.

Figure 9: Four regions in Botswana and the corresponding horizontal covariance function

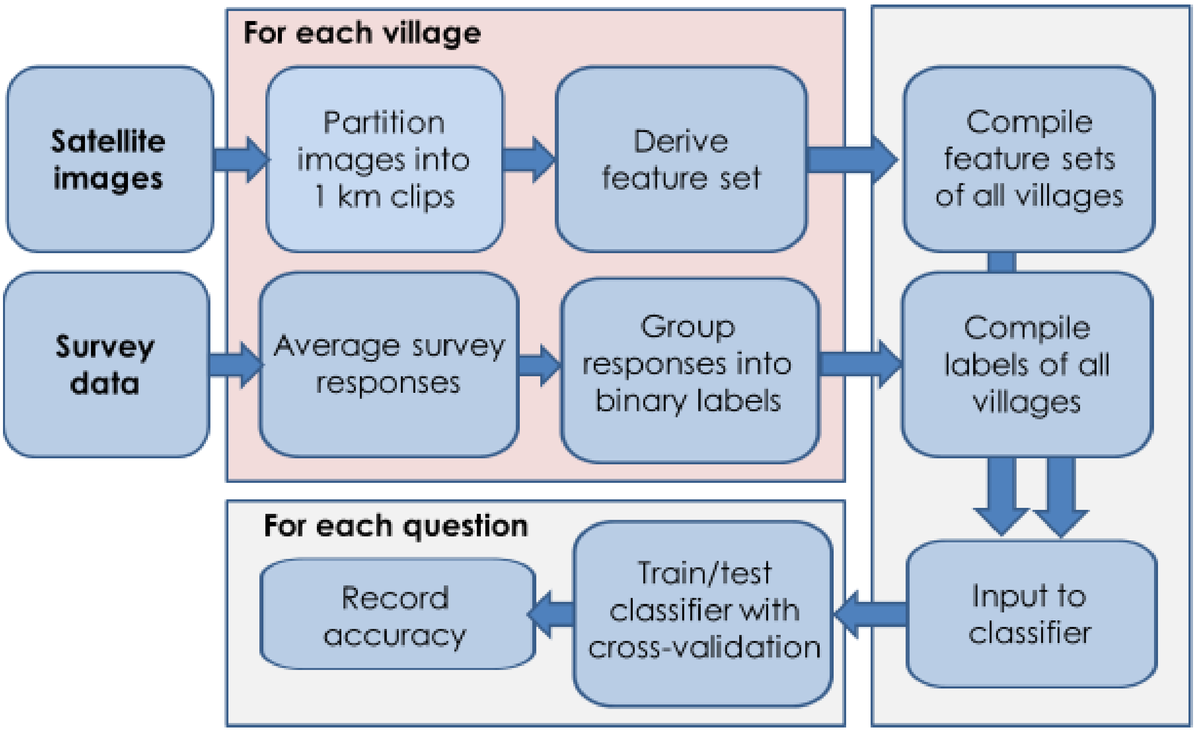


Imagery from DigitalGlobe provided by NGA

V.B Model development

In the model development process, we explored the features derived from the imagery to identify ones that provide useful information about the survey responses. Using the training data, we first analyzed the imagery-derived features in an exploratory mode to assess the utility of the features. The second step used the training data and automated feature selection to develop models for accurately predicting the survey-based indicators at the local level. The final step was to apply the trained model to the sequestered testing data to assess performance (Figure 10). In the remainder of this section, we present illustrative results linking specific image features to selected information from the survey data. These results focus on Botswana and explore issues of physical infrastructure and economic well-being. The next section presents the model validation and testing results for a sample of the survey questions.

Figure 10: Overview of model development and testing procedure



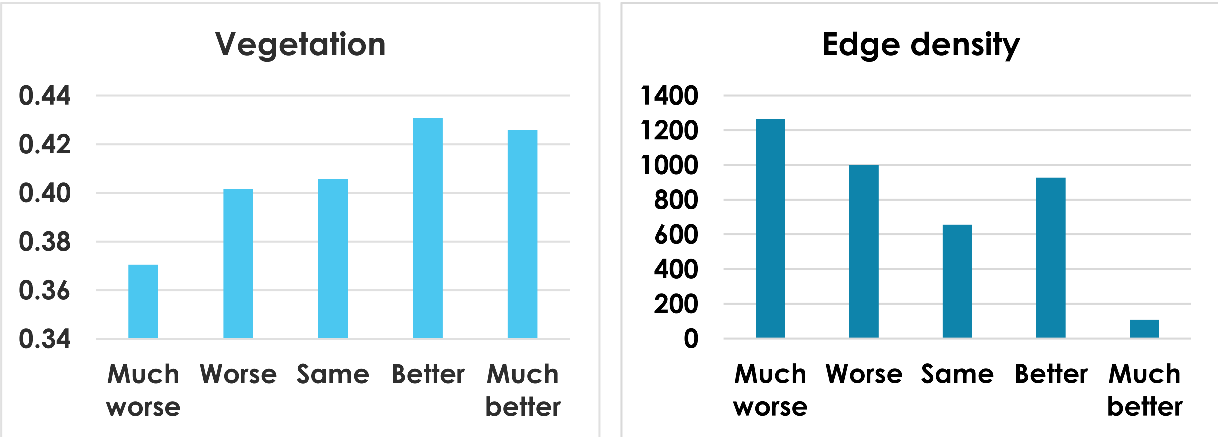
To illustrate the relationship between image features and survey responses, we present the mean responses for selected questions and the values of two image features. The first image feature is the vegetation measure, which is computed as the mean value across all pixels of the NDVI score. This provides a composite measure of vegetation and is sensitive to both the amount of vegetation in a location and the health of the vegetation. High scores are common in more affluent neighbourhoods. The second feature is the mean edge density in a region. As discussed above, edge density can vary non-linearly with socioeconomic status. Poor areas such as undocumented dwellings can have high edge density. However, middle-class neighbourhoods following planned layouts also have higher edge density than more spread-out suburban areas. We present the mean values of both image features relative to the responses to several survey questions: perceptions of one's own living conditions (Figure 11), access to electricity (Figure 12), access to piped water (Figure 13), access to cell-phone service (Figure 14), and access to schools (Figure 15). In all cases, the image features exhibit the pattern consistent with expectations. Poorer neighbourhoods have lower vegetation

Copyright ©Afrobarometer 2017

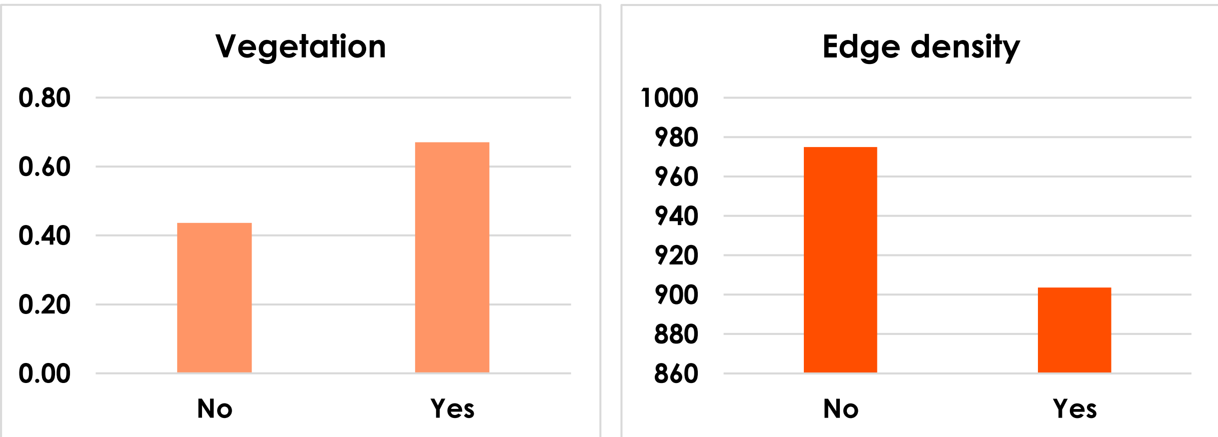
indices and higher edge densities. The non-linear behaviour of edge density is evident as well.

Figures 11-15: Comparison of average vegetation measure (NDVI) and average edge density to the mean responses in Botswana for the survey questions:

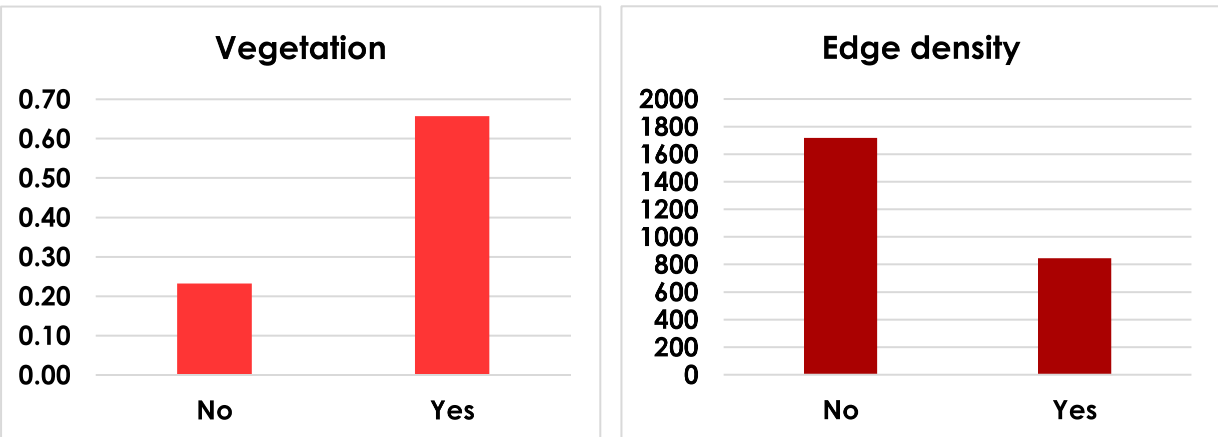
11: *"In general, how do you rate your living conditions compared to those of other Batswana?"*



12: *"Are the following services present in the primary sampling unit/enumeration area: Electricity grid that most houses could access?"*

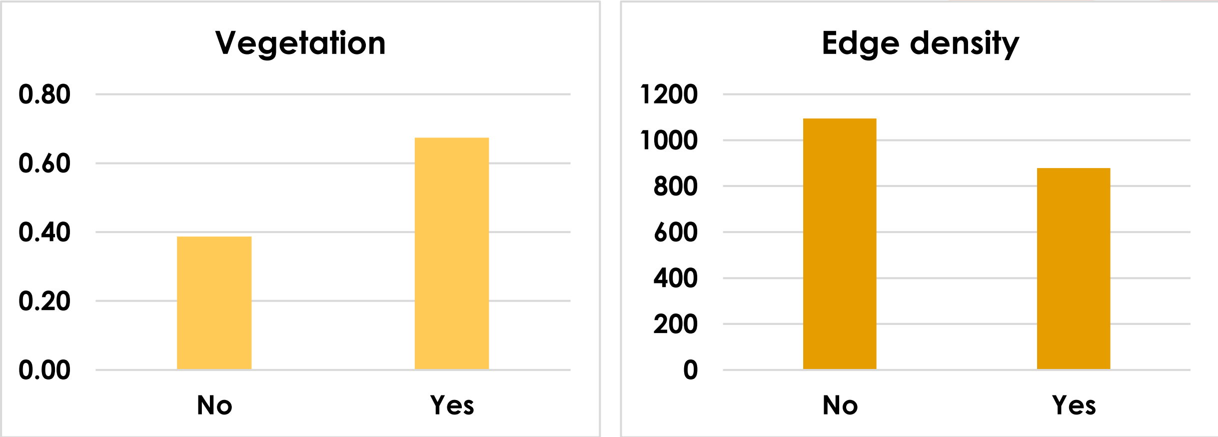


13: *"Are the following services present in the primary sampling unit/enumeration area: Piped water system that most houses could access?"*

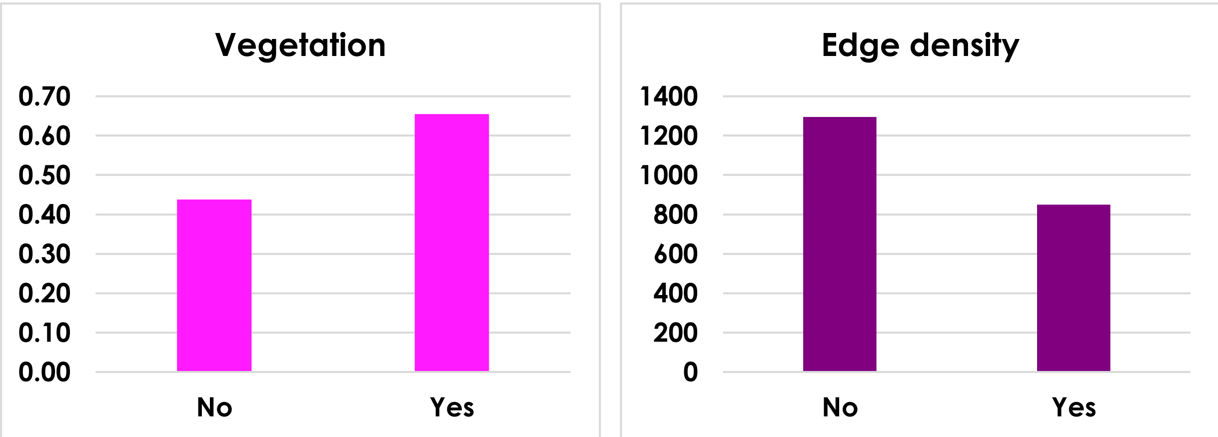


Copyright ©Afrobarometer 2017

14: *"Are the following services present in the primary sampling unit/enumeration area: Cell phone service?*



15: *"Are the following facilities available within the primary sampling unit/enumeration area, or within easy walking distance: Schools?"*



V.C. Model validation

To develop and validate the models, we randomly partitioned the data into training and test sets. Training of the classifier, including feature selection and model estimation, relied only on the training data. We scored the performance by comparing the image-based predictions to the values observed in the survey data, using only the testing data. The probability of correct classification computed from this analysis is our primary overall measure of the model's performance.

To demonstrate the process and illustrate performance, we present the results for selected survey questions using data from Botswana, Zimbabwe, and Kenya. The classifier was a k-nearest neighbour method with k = 3. The data were randomly partitioned such that 70% was for training and 30% was for testing. We explored two methods for modeling the survey data. One was to classify at the level of individual respondents, and the other was to aggregate the data to the village level and perform the classification for each location. Operating at the village level required assigning some type of consensus score to the survey responses for a single location. Conversely, because locations are known only to the village level, there was no way to use the imagery to differentiate individual responses within a single village.

Copyright ©Afrobarometer 2017

The model validation demonstrates the level of performance currently achieved using only image-based features to predict survey responses. Feature selection indicated that certain features are critical to these predictions: spatial features indicative of urbanization, spectral features quantifying vegetation health and prevalence, and indicators of social connectivity, such as road density and distance to major cities. Typical performance results appear in tables 5, 6, and 7. Table 5 depicts models for individual responses in Botswana, while Table 6 shows results aggregated to the village level. Results at the village level are generally more accurate. Turning to Zimbabwe, Table 7 shows performance for individual responses. The results for Kenya (Table 8) are far more preliminary. The only feature used for modeling for Kenya was a classification of the local area as either urban or rural. Even this crude level of information provides useful insights into access to infrastructure and quality of life. One interesting case, however, is access to cellular phone networks. Because Kenya has an extensive cellular network, knowing where a person lives (urban vs. rural) provides very little information. Effectively, almost everyone has cell phone access. Thus, the broad-based geospatial information derivable from commercial imagery does not provide useful information for predicting cell phone access. Higher-fidelity data that could identify specific cell towers might prove far more useful for this specific issue.

Table 5: Classifier performance for individual responses to selected questions in Botswana

Survey question	Valid responses	Percent correctly classified
House has access to electricity	Yes or no	91.9
House has access to piped water	Yes or no	94.1
House has access to sewage system	Yes or no	85.2
Economic conditions in the country	Good/neutral/bad	41.7
Own living conditions	Good/neutral/bad	45.6
Living conditions relative to other Batswana	5 categories	28.6
Access to enough food	Yes or no	57.1
Access to enough clean water	Yes or no	62.5
Fear of crime	Yes or no	58.5
Had something stolen from house	Yes or no	70.7

Table 6: Classifier performance for responses to selected questions in Botswana aggregated to the village level

Survey question	Valid responses	Percent correctly classified
Economic conditions in the country	Good/neutral/bad	55.6
Own living condition	Good/neutral/bad	83.3
Future economic outlook	Positive/negative	60.0
Access to enough food	Yes or no	68.2
Fear of crime	Yes or no	69.2
Had something stolen from house	Yes or no	57.9

Table 7: Classifier performance for individual responses to selected questions in Zimbabwe

Survey question	Valid responses	Percent correctly classified
House has access to electricity	Yes or no	95.5
House has access to piped water	Yes or no	93.2
Economic conditions in the country	Good/bad	59.2
Own living conditions relative to other	Good/bad	67.9
Had something stolen	Yes or no	58.3

Table 8: Classifier performance for individual responses to selected questions in Kenya

Survey question	Valid responses	Percent correctly classified
House has access to electricity	Yes or no	62.4
House has access to piped water	Yes or no	72.4
House has access to sewage system	Yes or no	89.2
House has access to cell phone network	Yes or no	28.0
Economic conditions in the country	2-class (bad/very bad vs. neutral/good/very good)	68.0

In addition to the illustrative results presented above, we explored the relative merits of different approaches to feature selection methods and different types of classifiers. For a set of 15 survey questions using the Botswana data, we applied various combinations of feature selection and classification methods and recorded the percent correctly classified on the testing set. The three feature selection methods were (1) use all available features, (2) select features based on joint mutual information (JMI), and (3) select features based on joint mutual information maximization (JMIM) (see Yang & Moody, 1999, and Bennasar, et al, 2015). The four classification methods were linear discriminant analysis (LDA), nearest neighbour, random forests, and support vector machine. The analysis showed negligible differences across feature selections methods. The choice of classifier, however, was important, with LDA performing significantly worse than the other three methods. This is not surprising, because LDA partitions the classes using hyperplanes on the implicit assumption of linear separability. The other three methods were generally comparable, although support vector machine had the best overall performance by a small margin.

VI. Findings

In the analysis of selected sub-Saharan African countries, image-derived features provide useful information for predicting survey responses across a range of questions. Performance is generally strongest on questions about infrastructure, such as access to electricity, clean water, and sewage disposal. Some questions about social attitudes, however, are performing only slightly better than chance. When compared to the results from our earlier study in Afghanistan, the performance here is less compelling.

There are several important differences between Afghanistan and sub-Saharan Africa that could account for the differences in performance. First, the locations in Afghanistan are

Copyright ©Afrobarometer 2017

almost universally small villages, whereas the study areas in Africa include larger, more dispersed built-up areas. Some of the Africa data include major cities such as Nairobi, Mombasa, Gaborone, and Harare. For these more extensive regions, the one-kilometer squares lack the spatial extent and context that could be very relevant to understanding local attitudes. A second factor is the difficulty in matching locations between imagery and survey data. For major urban areas, the initial data were insufficient to disambiguate locations within the city. We conducted further research into available gazetteers and were able to refine the location data, but this occurred late in the project. Using the improved information, a more focused analysis of urban areas is now possible. Finally, investigations into the specific image-derived features suggest that this is an area where significant improvements are possible, as discussed below.

VI.A. Image features and model refinement

We have identified several areas for future investigation that could improve model performance. There is a wide range of possible features that can be extracted from the imagery data. The initial exploration of the feature space was guided by the previous research in Afghanistan related to socioeconomic conditions. However, alternative features could provide better characterization of the phenomena, and there is considerable room for improvement.

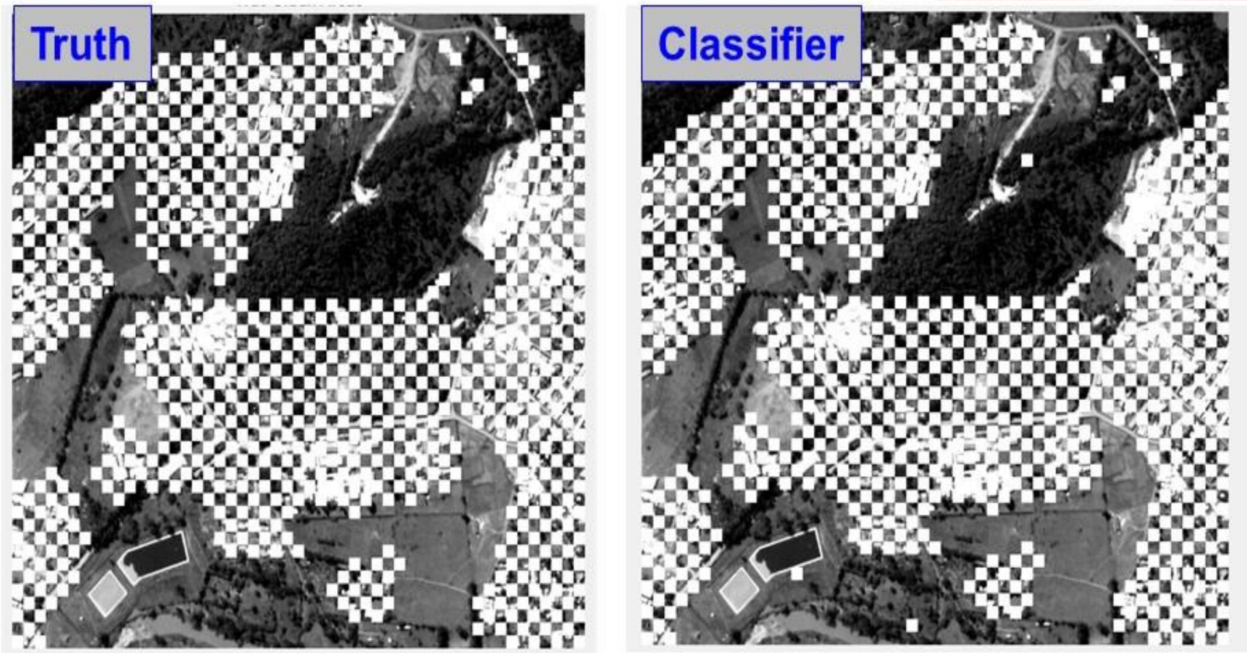
One class of features that merits exploration relies on a new method motivated by "bag of words" analysis of natural text. This new urban classification technique, called visual word region detection (VWRD), transforms an image into a set of visual words, and the statistics of the occurrence of each word are used as feature vectors. The words of urban and non-urban regions (as marked in training data) are put into histograms, and thus the most common and exclusive words for each region are used as indicators to segment and classify the test data (Weizman & Goldberger, 2009). One advantage of their technique is a normalization process that makes the method robust to changes in atmospheric conditions. We explored VWRD with some enhancements. To further improve the dictionary, we used scale-invariant feature transform (SIFT) to detect robust image features (Lowe, 2004). This differs from Weizman's VWRD, where principal component analysis (PCA) is used to transform the data patches in order to view and characterize the features of urban areas. The initial test of this method indicates good performance for classifying built-up areas (Figure 16). The next step is to test the method across a larger set of classes.

A second area for potential improvement in the models involves the characterization of the contextual information. As noted earlier, the one-kilometer squares used in this investigation provide no information about the surrounding conditions. The problem is especially evident in large urban areas, where the local neighbourhood significantly affects socioeconomic status. Figure 17 shows several neighbourhoods in Nairobi, and the differences in income level are evident from the imagery data.

We conducted some preliminary exploration to see whether machine-learning methods could provide an adequate statistical classification of the various urban areas. Our initial test employed a genetic algorithm approach developed for similar applications (Harvey et al., 2002; Harvey & Theiler, 2004). The initial classification results (Figure 18) offer a good starting point. However, additional post-processing could aid in refining the characterization of the urban areas. Morphological operators, for example, could be used to clean up discontinuities in the class overlay.

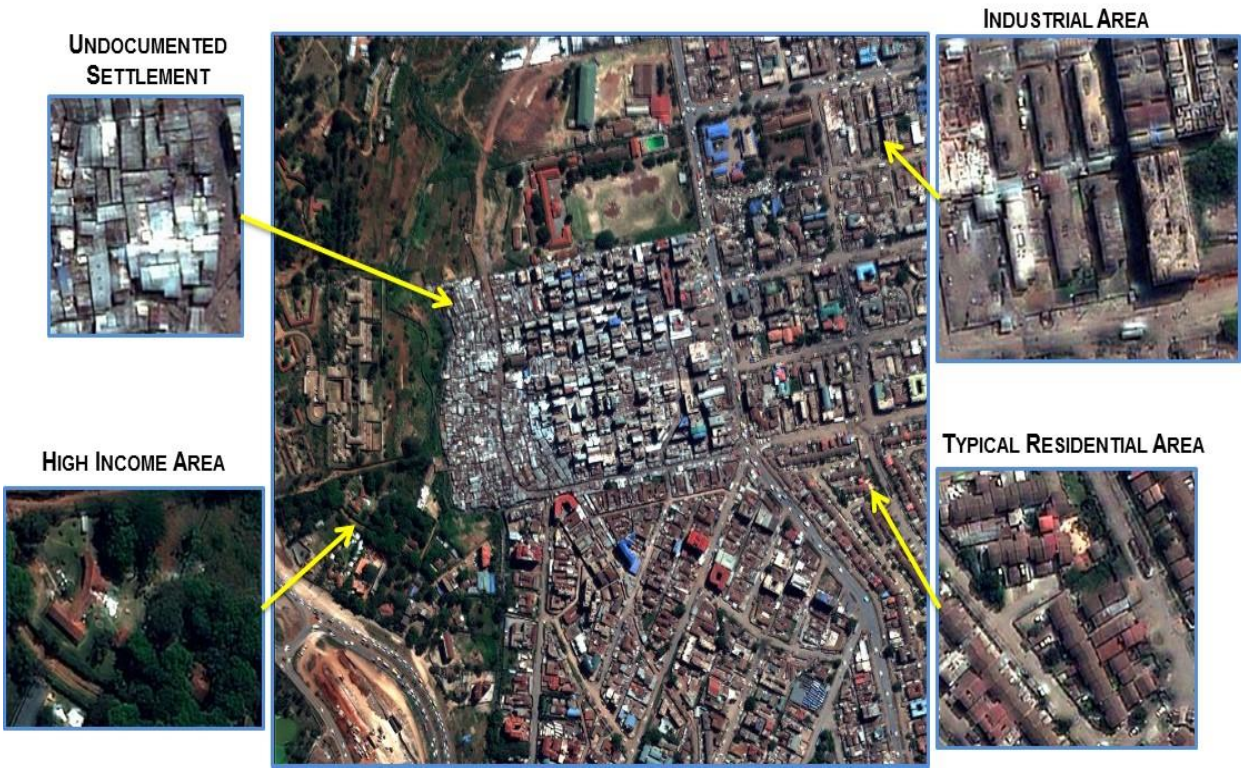
Copyright ©Afrobarometer 2017

Figure 16: Bag of visual words approach to classifying urban areas



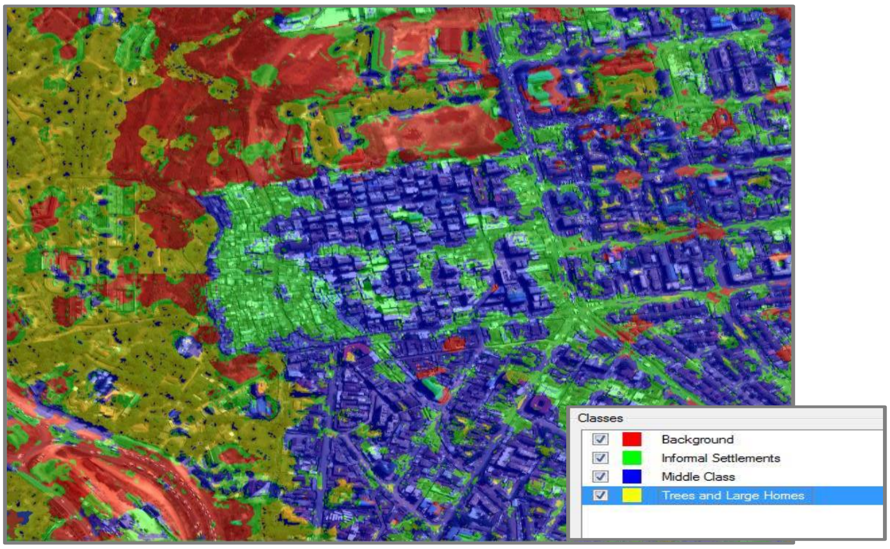
Imagery from DigitalGlobe provided by NGA

Figure 17: Diversity of land use and income level in urban areas



Imagery from DigitalGlobe provided by NGA

Figure 18: Initial urban area classification based on genetic algorithm



Imagery from DigitalGlobe provided by NGA

VI.B. Limitations with current models and potential temporal analysis

One limitation of the current work is that it captures a snapshot in time. Each survey occurs over a fairly short time interval, and the imagery was collected at dates close to the survey. For some locations, multiple images are available, and it is possible to assess changes over a short time period. One change that is evident in the imagery concerns vegetative health. Related to access to water, these changes can be due to seasonal variation or differences in weather patterns. In 2010 and 2011, for example, Kenya experienced its worst drought in 60 years (Mbogo, Inganga, & Maina, 2016). To illustrate the effects on image-derived features, Figure 19 presents two images of Harare, Zimbabwe, showing the differences in vegetation health.

Figure 19: Temporal differences in soil moisture and vegetation health due to seasonal patterns and drought conditions in Harare, Zimbabwe

Note: The image on the left, collected on 7 May 2010, immediately after the rain season, shows rich vegetation. The image on the right, collected on 2 September 2009, at the end of the dry period, shows the moisture-stressed vegetation. Imagery from DigitalGlobe provided by NGA

Copyright ©Afrobarometer 2017

Cloud cover presents another issue for image analysis. To address this, we performed a manual review of all images to identify and exclude images with cloud cover. A system for automatically excluding these images would be desirable. Even thin, wispy clouds pose a challenge (Figure 20). This is an area of active work in the remote sensing community.

Figure 20: Thin clouds obscure features of interest

Imagery from DigitalGlobe provided by NGA

VII. Conclusions and future research

The use of remote sensing to assess socioeconomic issues has numerous benefits, including lower risks and costs of data collection and data collection over extensive areas. Because commercial imagery data are readily available, measurements developed from these data can be readily shared without the restrictions associated with government-operated imaging systems. As companies such as Planet, Urthecast, and BlackSky deploy operational constellations, imagery will become more abundant, and finer-granularity temporal analysis will become possible.

This project focused on development and assessment of methods for deriving social, economic, and political indicators from imagery. The imagery portrays the physical attributes at a localized level. Recent advances in social science theory guided the investigation of physical features that are indicative of economic, political, and social attributes. By taking advantage of these relationships, this study sought to establish general methods for production of new types of useful geospatial intelligence information. Using survey data as the gold standard, we demonstrated methods for predicting these socioeconomic factors from the imagery data. Conversely, the analysis of physical observables in satellite imagery could in principle support independent verification of survey data. Because the satellite imaging can only observe physical phenomena, it cannot capture the feelings and thoughts of individuals. Thus, it is not a replacement for survey research, but rather an additional tool for understanding and characterizing society.

The performance of our models, while encouraging, has room for improvement. Several avenues of investigation are identified for refining and enhancing the models. As discussed in the previous section, alternative methods for extracting features from the imagery merit deeper investigation.

Using imagery data, the methods developed here and in related projects indicate that fundamental assessments of economic well-being, income distribution, and sociocultural factors can be derived from imagery. The imagery-derived information also serves as a basis for deeper analysis, providing the context for understanding more transient events. Monitoring changes over time can reveal trends in economic development, shifts in social organization, and possible early indicators of political changes. These indicators provide contextual information for fine-tuning the analysis of more transient data (e.g. news, social media) to yield better forecasting of major events. Two potential applications are prediction of elections and civil unrest.

The literature has suggested a link between economic development and election outcomes, but there remain a number of limitations, especially outside the U.S. context. Among the most prominent is the scale of measures for economic development. Studies suggesting that economic development, inequality, and public spending influence electoral outcomes are usually conducted on the national level, where data are more plentiful (e.g. Hibbs, 2012; Lewis-Beck & Tien, 2012). Local-level data are usually hard to access and compare across states. Moreover, local-level data are prone to bias. Available data on raw growth or government transfers to local areas can be confounded by inequality in the distribution of growth and by corruption. Satellite imagery provides us with unbiased measures of population growth, public investment, and inequality in localities. This is a potentially powerful tool for predicting district-level voting patterns in national elections and local-level election outcomes.

Overhead imagery data is too slow-moving to capture precipitating events, so it is unlikely to play the same role as social-media monitoring in predicting civil unrest. Overhead imagery can, however, provide valuable insight and context for the social-media data that are collected. While social media are good at capturing proximate factors related to civil unrest, such as the arrest of an opposition leader or an increase in public transportation fees, it is not as good at capturing underlying factors that lead to public dissatisfaction and likely influence the scale of unrest. The levels of political participation, public-goods provision, and social capital that are predictable from imagery can inform us about the degree to which a local population will take action. Overhead imagery can be utilized to capture some of these structural features, such as growing inequality, that underlie the scale of protest activity.

The need for future research is indicated, particularly with respect to persistent imagery of transient events and behaviours. With the emergence of small sats and the possibility of frequent imagery coverage, the range of social, economic, and political analysis supported by imagery data will grow enormously. In this study, as in the earlier study in Afghanistan, we essentially performed a static analysis of the society. Frequent imaging enables the observation of transient events and behaviours. Exploitation of this emerging source could support deeper analysis of political issues, such as elections and civil unrest. Furthermore, it offers a valuable data source for understanding disease outbreak and spread (Ramakrishnan, 2014), intermodal transportation, and industrial production (Irvine, 2015).

References

Acemoglu, D., & Robinson, J. (2009). *Economic origins of dictatorship and democracy*. Cambridge, UK: Cambridge University Press.

Alparone, L.,Wald, L., Chanussot, J.,Bruce, L. M., & Thomas, C. (2007). Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest. *IEEE Transactions on Geoscience and Remote Sensing*, *45*(10), 3012-3021.

Bagus, P. (2006). The problem of accuracy of economic data. Mises Institute. https://mises.org/library/problem-accuracy-economic-data.

Bardhan, P., Bowles, S., & Gintis, H. (2000). Wealth inequality, credit constraints, and economic performance. In A. Atkinson & F. Bourguignon (Eds.), *Handbook of Income Distribution.* Dortrecht: North-Holland.

Bennasar, M., Hicks, Y., & Setchi, R. (2015). Feature selection using joint mutual information maximisation. *Expert Systems With Applications*, 42, 8520-8532.

Bhaduri, B., Bright, E., Coleman, P., & Urban, M. (2007). LandScan USA: A high resolution geospatial and temporal modeling approach for population distribution and dynamics. *GeoJournal, 69*, 103-117.

Blumberg, D., & Jacobson, D. (1997). New frontiers: Remote sensing in social science research. *American Sociologist, 28*(3), 62-68.

Bowles, S., & Gintis, H. (2002). Social capital and community governance. *Economic Journal*, *112*, 419-436.

Boix, C. (2003). *Democracy and redistribution*. Cambridge, UK: Cambridge University Press.

Cheriyadat, A., Bright, E. A., Bhaduri, B., & Potere, D. (2007). Mapping of settlements in high resolution satellite imagery using high performance computing. *GeoJournal, 69,*119-129.

Crews, K. A., & Walsh, S. J. (2009). Remote sensing and the social sciences. In T. Warner, D. Nellis, and G. Foody (Eds.), *Handbook of Remote Sensing*, pp. 437-435. Sage Publications.

Elvidge C. D., Baugh, K. E., Kihn, E. A., Kroehl, H. W., Davis, E. R., & Davis, C. W. (1997). Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption. *International Journal of Remote Sensing, 18*, 1373-1379.

Fox, J., Rindfuss, R., Walsh, S. J., & Mishra, V. (eds). (2003). *People and the environment: Approaches for linking household and community surveys to remote sensing and GIS*. Boston: Kluwer.

Goodchild, M. F., Anselin, L., Appelbaum, R. P., & Harthorn, B. H. (2000). Toward spatially integrated social science. *International Regional Science Review, 23*(2), 139-159.

Habyarimana, J., Humphreys, M., Posner, D., & Weinstein, J. (2007). Why does ethnic diversity undermine public goods provision? *American Political Science Review, 101*(4), 709-725.

Hall, O. (2010). Remote sensing in social science research. *Open Remote Sensing Journal*, *3*, 1-16, 1875-4139, Bentham Open.

Harvey, N. R., Theiler, J., Brumby, S. P., Perkins, S., Szymanski, J. J., Block, J. J., Porter, R. B., Galassi, M., & Young, A. C. (2002). Comparison of GENIE and conventional supervised classifiers for multispectral image feature extraction. *IEEE Transactions on Geoscience and Remote Sensing, 40,* 393-404.

Harvey, N. R., & Theiler, J. (2004). Focus-of-attention strategies for finding discrete objects in multispectral imagery. *Proceedings of SPIE, 5546,* 179-189.

Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review, 102*(2), 994-1028.

Hibbs, D. A. (2012). Obama's reelection prospects under "bread and peace" voting in the 2012 US presidential election. *PS: Political Science & Politics, 45*(4), 635-639.

Irvine, J. M. (2003). National imagery intelligence rating scale (NIIRS). In M. Dekker (Ed.), *The Encyclopedia of Optical Engineering*.

Irvine, J. M., Lepanto, J., Regan, J., & Young, M. (2012). Deriving economic and social indicators from imagery. IEEE Applied Imagery and Pattern Recognition Workshop. Washington, DC: Cosmos Club.

Irvine, J., Lepanto, J., Markuzon, N., Regan, J., Vaisman, E., Young, M., Christia, F., & Petersen, R. (2013). Characterization of villages in rural Afghanistan. Proceedings from American Society of Photogrammetry and Remote Sensing (ASPRS) 2013 Annual Conference, Baltimore, MD.

Irvine, J., & Lepanto, J. (2014). Remote sensing and indicators of well-being and governance. Draper Technical Report.

Irvine, J. M. (2015). Large scale analytics in the era of abundant data. ENVI Analytics Symposium (EAS), 25-26 August 2015, Boulder, CO.

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science, 353*(6301).

Jenson, J. R., & Cowen, D. C. (1999). Remote sensing of urban infrastructure and socio-economic attributes. *Photogrammetric Engineering and Remote Sensing*, *65*(5), 611-622.

Jerven, M. (2014). On the accuracy of trade and GDP statistics in Africa: Errors of commission and omission. *Journal of African Trade, 1*(1) 45-52. https://doi.org/10.1016/j.joat.2014.08.001.

Jiang, H. (2003). Stories remote sensing images can tell: Integrating remote sensing analysis with ethnographic research in the study of cultural landscapes. *Human Ecology*, *31*(2), 215-232.

Kennedy, R. (2010). The contradiction of modernization: A conditional model of endogenous democratization. *Journal of Politics, 72*(3), 785-798.

Knack, S., & Keefer, P. (1997). Does social capital have an economic payoff? A cross-country investigation. *Quarterly Journal of Economics, 112*(4), 1251-1288.

Kriegler, F. J., Malila, W. A., Nalepka, R. F., & Richardson, W. (1969). Preprocessing transformations and their effects on multispectral recognition. *Proceedings of the Sixth International Symposium on Remote Sensing of Environment*, University of Michigan, Ann Arbor, MI, 97-131.

Kruse, F. A., Lefkoff, A. B., Boardman, J. B., Heidebrecht, K. B., Shapiro, A. T., Barloon, P. J., & Goetz, A. F. H. (1993). The spectral image processing system (SIPS): Interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, *44*, 145-163.

Landscan Project at Oak Ridge National Lab: http://www.ornl.gov/sci/landscan/.

Lee, J., & Lee, C. (2010). Fast and efficient panchromatic sharpening. *IEEE Transactions on Geoscience and Remote Sensing*, *48*(1).

Lewis-Beck, M. S., & Tien, C. (2012). Election forecasting for turbulent times. *PS: Political Science & Politics*, 45(4), 625-629.

Leachtenauer, J.C., Malila, W., Irvine, J. M., Colburn, L., & Salvaggio, N. L. (1997). General image quality equation. *Applied Optics*, *36*(32), 8322-8328.

Leachtenauer, J.C., & Driggers, R.G. (2001). *Surveillance and reconnaissance systems: Modeling and performance prediction.* Artech House.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision, 60*(2), 91-110.

Copyright ©Afrobarometer 2017

Mbogo, E., Inganga, F., & Maina, J. M. (2016). Drought conditions and management strategies in Kenya. http://www.ais.unwater.org/ais/pluginfile.php/601/mod_page/content/29/Kenya.pdf.

Min, B., Agnew, J., Gillespie, T. W., & Gonzalez, J. (2008). Baghdad nights: Evaluating the US military surge using nighttime light signatures. *Environment and Planning A, 40*(10), 2285-2295.

National Research Council. (1998). *People and pixels: Linking remote sensing and social science.* Commmittee on the Human Dimensions of Global Change.

O'Brien, M. A., & Irvine, J. M. (2004) Information fusion for feature extraction and the development of geospatial information. International Conference on Information Fusion, Stockholm, Sweden.

Owen, K. K. (2012). Geospatial and remote sensing-based indicators of settlement type – differentiating informal and formal settlements in Guatemala City. Ph.D. dissertation, George Mason University.

Ramakrishnan, N., et al. (2014). 'Beating the news' with EMBERS: Forecasting civil unrest using open source indicators. Proceedings of the 20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (Industrial Track), August 2014.

Rashmi, S., Addamani, S., Venkat, & Ravikiran S. (2014). Spectral angle mapper algorithm for remote sensing image classification. *International Journal of Innovative Science, Engineering & Technology*, *1*(4).

Rauk, J. (1993). Productivity gains from geographic concentration of human capital. *Journal of Urban Economics, 34*(3), 380-400.

Rodrik, D. (2000). Institutions for high-quality growth. *Studies in Comparative International Development, 35*(3), 3-31.

Taubenböck, H., Wurm, M., Setiadi, N., Gebert, N., Roth, A., Strunz, G., Birkmann, J., & Dech, S. (2009). Integrating remote sensing and social science. *Joint Urban Remote Sensing Event*, 20-22 May 2009, Shanghai, China.

Vijayaraj V., Bright E. A., & Bhaduri, B. L. (2007). High resolution urban feature extraction for global population mapping using high performance computing. Proceedings of 2007 IEEE International Geosciences and Remote Sensing Symposium.

Vijayaraj V., Cheriyadat, A. M., Sallee, P., Colder, B., Vatsavai, R. R., Bright E. A., & Bhaduri B.L. (2008). Overhead image Ssatistics. Proceedings of 2008 IEEE 37th Applied Imagery Pattern Recognition Workshop.

Weizman, L., & Goldberger, J. (2009). Detection of urban zones in satellite images using visual words. *IEEE Geoscience and Remote Sensing Letters*, *6*(3).

World Bank. (2017). Living Standards Measurement Study. Available at:

http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTLSMS/0,,contentMDK:21485765~menuPK:4196952~pagePK:64168445~piPK:64168309~theSitePK:3358997~isCURL:Y,00.html.

Yang, H. H., & Moody, J. (1999). Feature selection based on joint mutual information. Advances in Intelligent Data Analysis (AIDA), Computational Intelligence Methods and Applications (CIMA), International Computer Science conventions, Rochester, NY.

Recent Afrobarometer working papers

No. 177 Depetris-Chauvin, Emilio & Ruben Durante. One team, one nation: Football, ethnic identity, and conflict in Africa. 2017.

No. 176 Tannenberg, Marcus. The autocratic trust bias: Politically sensitive survey items and self-censorship. 2017.

No. 175 Liu, Shelley. Wartime educational loss and attitudes toward democratic institutions. 2017.

No. 174 Crisman, Benjamin. Disease, disaster, and disengagement: Ebola and political participation in Sierra Leone. 2017.

No. 173 Claassen, Christopher. Explaining South African xenophobia. 2017.

No. 172 Logan, Carolyn. 800 languages and counting: Lessons from survey research across a linguistically diverse continent. 2017.

No. 171 Guardado, Jenny & Leonard Wantchekon. Do electoral handouts affect voting behavior? 2017.

No. 170 Kerr, Nicholas & Anna Lührmann. Public trust in elections: The role of media freedom and election management autonomy. 2017.

No. 169 McNamee, Lachlan. Indirect colonial rule and the political salience of ethnicity. 2016.

No. 168 Coulibaly, Massa. Measuring democracy in Africa: Applying anchors. (French). 2016.

No. 167 Monyake, Moletsane. Does personal experience of bribery explain protest participation in Africa? 2016.

No. 166 Robinson, Amanda Lea. Ethnic diversity, segregation, and ethnocentric trust in Africa. 2016.

No. 165 Hounsounon, Damas. Décentralisation et qualité de l'offre de services socio-publics en Afrique subsaharienne. 2016.

No. 164 Mattes, Robert & Mulu Teka. Ethiopians' views of democratic government: Fear, ignorance, or unique understanding of democracy? 2016.

No. 163 Manacorda, Marco & Andrea Tesei. Liberation technology: Mobile phones and political mobilization in Africa. 2016.

No. 162 Albaugh, Ericka A. Language, education, and citizenship in Africa. 2016.

No. 161 De Kadt, D., & Evan S. Lieberman. Do citizens reward good service? Voter responses to basic service provision in southern Africa. 2015.

No. 160 Aquino, J. A. The effect of exposure to political institutions and economic events on demand for democracy in Africa. 2015.

No. 159 Tsubura, Machiko. Does clientelism help Tanzanian MPs establish long-term electoral support? 2015.

No. 158 Claire, Adida L., Karen E. Ferree, Daniel N. Posner, & Amanda L. Robinson. Who's asking? Interviewer coethnicity effects in African survey data. 2015.

No. 157 Bratton, Michael & E. Gyimah-Boadi. Political risks facing African democracies: Evidence from Afrobarometer. 2015.

No. 156 Croke, Kevin, Guy Grossman, Horacio A. Larreguy, & John Marshall. Deliberate disengagement: How education decreases political participation in electoral authoritarian regimes. 2015.

Afrobarometer Working Papers Series

Editor: Michael Bratton, mbratton@msu.edu
Editorial Board: E. Gyimah-Boadi, Carolyn Logan, Robert Mattes, Leonard Wantchekon

Afrobarometer publications report the results of national sample surveys on the attitudes of citizens in selected African countries toward democracy, markets, civil society, and other aspects of development. Afrobarometer publications are simultaneously co-published by the six Afrobarometer Core Partner and Support Unit Institutions. All Afrobarometer publications can be searched and downloaded from www.afrobarometer.org.

Support for Afrobarometer's current Round 7 surveys is provided by the Swedish International Development Cooperation Agency (SIDA), Mo Ibrahim Foundation, Open Society Foundations, Bill & Melinda Gates Foundation, William and Flora Hewlett Foundation, U.S. State Department, National Endowment for Democracy, and Transparency International.

Core partners:

Center for Democratic Development (CDD-Ghana)
95 Nortei Ababio Street, North Airport Residential Area
P.O. Box LG 404, Legon-Accra, Ghana
Tel: +233 21 776 142
Fax: +233 21 763 028
www.cddghana.org

Institute for Development Studies (IDS), University of Nairobi
P.O. Box 30197, Nairobi, 00100, Kenya
Tel: +254 20 2247968
Fax: +254 20 2222036
www.ids.uonbi.ac.ke

Institute for Empirical Research in Political Economy (IREEP)
Arconville, Lot 104 - Parcelle J, 02 BP: 372, Cotonou, Republique du Benin
Tel: +229 21 363 873/ 229 94 940 108
Fax: +229 21 362 029
www.ireep.org

Institute for Justice and Reconciliation (IJR)
105 Hatfield Street, Gardens, 8001, Cape Town, South Africa
Tel: +27 21 763 7128
Fax: +27 21 763 7138
www.ijr.org.za

Support units:

Michigan State University (MSU)
Department of Political Science
East Lansing, MI 48824, USA
Tel: +1 517 353 6590; Fax: +1 517 432 1091
www.polisci.msu.edu

University of Cape Town (UCT)
Institute for Democracy, Citizenship and Public Policy in Africa
Leslie Social Science Building
Rondebosch, Cape Town, WC 7701
South Africa
Tel: +27 21 650 3827