SETTLEMENT INDICATORS OF WELLBEING AND ECONOMIC STATUS – LACUNARITY AND VEGETATION

Karen Owen, PhD Candidate
George Mason University
Department of Geography and Geoinformation Science
Fairfax, VA 22030
kowen@gmu.edu

ABSTRACT

This research combines geospatial and image analysis to develop indicators of wellbeing that distinguish poor/informal communities from wealthy/planned ones in a study site in Latin America. By masking to the irregular-shaped neighborhoods of dwelling-areas enclosed by roads and terrain features, we zero-in on those characteristics that apply to neighborhoods where people live and not the surrounding, unrelated features. Using the actual shapes of settled areas and their spectral values from high resolution satellite imagery we can elicit key structural components of vegetation and the built-up in-between gaps to develop indicators that allow us to compare and differentiate settlement types. The objective is to quantify a settlement type using applied measures of lacunarity, fractal dimension, and landscape patch attributes to evaluate relative poverty or wealth, and build a better understanding of the quality of life of the underlying residents.

KEYWORDS: Lacunarity, fractal dimension, informal settlements, quality of life, vegetation, human geography

INTRODUCTION AND PROBLEM STATEMENT

This research has been conducted in search of applied methods to evaluate differences in settlement structure based on theoretical understanding of the expected differences in shape and texture of formal (planned), and informal (unplanned) communities. This is the first study of its kind to apply such research to neighborhoods in Meso-America, specifically Guatemala City. Reliable measures are needed that can be applied to masked settlement areas instead of convenient rectangular matrices, as prior efforts have done. This research tests a series of methods to the current study area using standard Quickbird imagery to re-produce the methods found in the literature. This will assess their combined usefulness in quantifying differences between settlements as proxies for quality of life and economic status largely derived from imagery. In this case, our prior knowledge of the actual settlement’s economic status was used to test the measures. Future work would apply the same measures without reference data, to determine in which class the settlement belongs.

PRIOR RESEARCH

Fractal Dimension of Settlements

Evaluating the fractal dimension of settlement areas, both formal and informal, is considered a scale-invariant technique (Thomas et al. 2008; Tannier & Pumain 2005; Cooper 2005). Wallace et al. (2004) studied the fractal dimension of built up areas at varying scales to confirm this. As described by Tannier and Pumain (2005) “the main advantage of fractal geometry is to provide a model of reference which seems more adapted than Euclidean geometry to the description of spatial forms created by societies: features of heterogeneity, self-similarity and hierarchy are included from the very beginning in fractal structures”. Research to assess spatial heterogeneity of pattern in urban areas was conducted by Cooper (2001) who evaluated the fractal pattern of street edges in relation to their building sizes and locations, finding that higher fractal dimensional values were found in areas of increasing fragmentation of the building line as separation differences between building sizes increased. Fractal dimension also decreased as building size increased (Ibid, 103). Fractal dimension of housing patches in different sprawl neighborhoods was evaluated by LaGro (1998) as a method to understand neighborhood housing context based on neighborhood spatial placement. Fractal analysis has been conducted at the city level where fractal dimensions approaching 2 represent a compact mass, and those approaching 1 are linear (Batty 2009). As Thomas et al. (2008) specify, “fractal measures unequivocally characterise the spatial organisation of urban patterns. They can be used to...
measure the extent to which built-up areas are distributed in a uniform or spasmodic way”. If this is true, then fractal dimension should yield significant differences between types of settlements (rich vs. poor; planned vs. unplanned).

**Lacunarity of Settlements**

Another theoretical measure found in the informal settlement literature is lacunarity, represented as $\Lambda$. Simply stated, lacunarity measures “gappiness” or “visual texture”, and is a measure of translational or rotational invariance or heterogeneity in an image. Lacunarity supplements fractal dimension to characterize settlement patterns and has been used to evaluate informal settlement housing (Filho & Sobreira 2005, 2007, 2008; Junior & Filho 2005) to compare neighborhood racial segregation at multiple scales (Wu et al. 2001), and to determine whether the measure helps improve classification accuracy (Myint et al. 2006). After converting an image to binary so that dwellings or objects the analyst wishes to focus on become foreground pixels, lacunarity should identify heterogeneity as well as regularity. In general terms, lacunarity represents the variation in foreground pixel density over various box sizes moving over the image window.

The most-studied lacunarity algorithms include the Gliding Box algorithm and the Differential Box-Counting (DBC) algorithm. Originally, the gliding box algorithm was recommended over the box-counting method in areas with limited data samples when used as a multifractal method for geochemical data (Cheng 1999). Later, the DBC algorithm was found to correctly classify 90% of image sub-scenes in the region of Recife, Brazil as informal settlements, versus 80% correctly identified using the Gliding Box algorithm (Filho & Sobreira 2008). In other work, an “inhabitability index” for settlements in Brazil was developed that concluded the DBC algorithm applied to binarized grayscale Quickbird images was best able to discriminate texture in urban areas of differing inhabitability conditions (Filho & Sobreira 2007). It was found that slum areas exhibit lower lacunarity values resulting from stated “lower permeability” and lower “gappiness” (Ibid). In this case, the authors assigned an inhabitability index based on socioeconomic values of a census region’s central pixel (aggregated to the geographic coordinates of its regional capital), then developed a kriged surface. This artificially assumed that economic status as a measure of inhabitability can be derived from the aggregation of census data to a reporting point and is inversely proportional (to a scaling factor) as distance away from the central pixel increases, but that inhabitability is also related by proximity to a nearby region capital’s socioeconomic values. Small samples were then selected randomly and assigned high or low inhabitability based on the kriged values they overlaid. The samples were histogram-equalized, which alters the original pixel values, and then converted to binary using an unspecified threshold. The authors did not classify the built and non-built environments, which could potentially have an impact on ability to delineate urban features. Although this novel method is interesting in its use of lacunarity to evaluate urban texture, the inhabitability index (as a proxy for socio-economic status) was estimated somewhat artificially. Lacunarity values have also been shown to vary by scale as well as settlement type using extracted building shapes in rectangular-shaped image subsets (Junior & Filho 1997). However, it was also shown that more regularized settlements representing the city center and not the squatter/informal settlements, had higher lacunarity values at similar scales and not lower values (Ibid, 12). Contrastingly, Myint et al. (2006) found that higher spatial heterogeneity, as one would expect in unplanned communities, resulted in higher lacunarity values. Given the disagreement in results, we sought to determine whether lacunarity could help distinguish differences in settlement types in Guatemala City, and at what spatial scale.

**Vegetation and Greenspace of Settlements**

Niebergall et al. (2007) developed a complex processing chain to extract informal settlement areas from Delhi, India using Quickbird VHR imagery with the objective of rating quality of socio-economic indicators of population to ultimately estimate water consumption and waste-water disposal needs. An intensive field campaign was conducted to sample and identify housing and settlement types, water-related structures and distinct features in situ that could be identified in the Quickbird scenes. In addition, surveys gathered family size and water consumption data. The authors sought to test the integration of segmented object extraction combined with GIS to evaluate vulnerability of informal settlements within mega cities and to determine if their methods could derive the socio-economic indicator values indirectly via image analysis. The authors’ test sites were selected for their high socio-economic gradient-settlement structures from upper and middle class residential areas were near informal settlements. Quickbird imagery was pan-sharpened and a semi-automated supervised classification of residential structures (imperviousness), roads (soil), vegetation, and shadows was conducted. The authors found that texture was an essential parameter in the detection of informal settlements, and that classification at varying scales enabled them to extract small features such as houses and larger features such as streets. From some level of trial and error they determined the appropriate classification level for each feature, calibrated the rule set on the training site, and

**Pecora 18 – Forty Years of Earth Observation...Understanding a Changing World**

**November 14 – 17, 2011 * Herndon, Virginia**
integrated the survey results with the imagery analysis to estimate urban vulnerability to unmet water resource needs.

This research is particularly significant because it sought to compare both informal and formal settlements in the same region, and to determine if socio-economic indicators can be measured from imagery. The use of accepted measures such as house size in $m^2$, imperviousness, and Normalized Difference Vegetation Index (NDVI) underscores the importance of these remote sensing indicators to understand economic status. One of many indicators hypothesized to impact quality of life metrics was vegetation health (Cowen & Jensen 2001:176). The impact of informal settlements on the surrounding environment has received some interest in measuring degree of environmental degradation caused by slum growth. Zeilhofer and Topanotti (2008) used a variety of variables quantified through prior research on secondary data, field work, air photo interpretation and GIS techniques through redundancy analysis to evaluate the environmental impacts. The use of detailed municipal planning data enabled the development of very specific indicators. Indicators related to slum housing included distance to street, amount of greenspace, and sidewalks. The Zeilhofer and Topanotti (2008) work could only be repeated with extremely detailed urban planning data available for validation, but this recent work emphasizes the importance of including green space in measuring informal settlement variation.

Borrowing from landscape patch dynamics, settlements have been modeled using algorithms well-developed for landscape analysis: area, density, form, edge, core area, proximity, subdivision and diversity (Lang et al. 2009; Sudhira et al. 2004; Yeh & Li 2001). Entropy, patchiness, and density/growth of built-up areas was measured by Sudhira et al. (2004). Although their study applied previously developed metrics for landscape analysis to the sprawl growth in Mangalore, India, it did not focus specifically on informally-settled or slum areas (Ibid). In their research, similar buildings were segmented, classified and clumped into a homogenous landscape ‘patch’ and then patch types were classified using maximum-likelihood into the five classes of built-up, vegetation, water, agricultural land and open land (Ibid).

The prior literature devoted to measuring texture differences in the spatial structure of settlements suffers from several drawbacks. Lacunarity shows both high and low values for informal-type settlements, which brings into question its utility as a metric that can be applied to understand typologies. Fractal dimension has generally been measured at the city or regional level, and there are questions regarding what features should become the foreground to be measured and why. Since vegetation is a known quality of life indicator, patch characteristics will be used in this research to understand settlement differences. The variety of classification methods, threshold values for urban scenes, and the reliance on rectangular shapes is also limiting. Settlements do not exhibit rectangular shapes – the neighborhoods humans inhabit are normally constrained by topography, placement of pre-existing major roads, and zoning laws, and therefore nearly always contain irregular boundaries.

**STUDY AREA**

The study area (Figure 1) is in Zona 11 in Guatemala City, Guatemala, west of La Aurora International airport. Twelve distinct settlements were defined a-priori from expert knowledge and field work, including driving through...
and viewing most of the settlements (Partner for Surgery, 2009-2010). Entry to some was barred by armed guards and others had to be viewed from a distance due to safety concerns. Settlement boundaries were estimated through a combination of Google Earth visual interpretation aided by half-meter orthoimage comparison and discussions with local experts, and by consulting community names on a topographic map. The following qualitative guidelines constrained the settlement selections:

- Exclude larger buildings adjacent to major roads – many are commercial areas, not dwellings
- Exclude surrounding highways – generally not considered part of the neighborhood
- Exclude wooded or vegetated areas outside the neighborhood – does not characterize dwelling areas

The settlements were created from a 36m² Quickbird image scene recorded in March 2009 (Digital Globe, 2009).

METHODS

Three types of settlement metrics were computed: fractal dimension, lacunarity, and vegetation-related. A major goal of the research was to evaluate settlement separability from each measure as compared to what was found in the literature. If we understand what distinguishes between planned (rich) and unplanned (poor) settlements, then we should have proxy variables for economic status and wellbeing. These measures were not expected to exhibit multicollinearity being computed separately, so we did not evaluate covariance from a regression standpoint. For each measure, the two-sample unpaired t-test of means was calculated to determine probability of each group originating from different populations. The unpaired t-test considering group A (formal) and group B (informal) to evaluate statistical significance is:

\[ t = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{\frac{\text{Var}(x_A)/N_A + \text{Var}(x_B)/N_B}} \}

We are reminded that very small differences in measured results can still be significant when the sample size (degrees of freedom) is large. Each image subset was digitally masked to its non-rectangular boundary so that exterior pixels and shapes containing unrelated settlement features were not included in the computations. This method differs from most prior settlement studies that compute every measure using a rectangular shape. Prior work either includes unrelated external features to facilitate ease of matrix computation or excludes edge regions that are still part of the settlement but are discarded to create a more ‘pure’ interior image scene. Assuming the morphology of a neighborhood is rarely if ever a rectangle, the bias that could potentially result from over-including or over-excluding portions of the total settled area was avoided.

Pre-processing included computing the normalized difference vegetation index (NDVI) from the Red and NIR Quickbird bands for each masked region using a threshold value of 0.09 in order to create vectors of each vegetated patch, producing a reasonable approximation of vegetation within each urban area. The image processing tool ENVI (ITT-VIS, 2009) was used to orthorectify and pansharpen the image scene, perform masking, compute vegetation indices, threshold and then vectorize the results for the vegetation measures. For lacunarity, the grayscale images were rescaled to eliminate negative NDVI results which would produce undefined lacunarity values. Using this model, non-vegetation was treated as pseudo built-up, and then binarized such that foreground pixels represented non-vegetation while background pixels were vegetation. The resulting binary and its precursor grayscale image were used for lacunarity and fractal image analysis (Rasband, 2010). For the lacunarity and the fractal calculation, this author considered anything not comprised of vegetation (roads, dwellings, buildings, etc.) in highly urban areas can be considered built-up evidence of human habitation. Due to fairly small size of masked image subsets (m masked area was 0.3km²) the differential box counting method was run using experimental box sizes of 25, 100, 200, and 500 to determine which scale performed best at discriminating settlement type, and therefore its economic status. The box size of 200 performed best, so only those results are displayed.

RESULTS

Fractal Dimension

Fractal dimension is calculated using the formula: \( D_B = -\lim[\log N_e/\log \varepsilon] \) which is read as the negative limit of the ratio of the log of the number of boxes at a certain scale over the log of that scale (Karperian, N.D). The results of calculating fractal dimension of the binary image using the sliding box method and box size (scale) of 200 produced
the only promising result - a 92% confidence level that fractal dimension of built-up areas in the informal vs. the planned communities yield different results but, with no significant difference detected on the grayscale image, with \( p(0.58) \). As an interesting exercise and because of the simplicity of creating a binary representation of all asphalt roads following classification, \( D_B \) was also computed on asphalt roads in the image subsets, yielding no significant difference. This is likely because the same underlying structural process contributed to their creation regardless of whether the settlement was rich or poor, and the same government infrastructure resources would have been used to produce them. We can conclude that despite the additional detail afforded by variations in pixel values in the grayscale images, this detail is too difficult to distinguish discrete structural content of the settlement’s built-up areas, and was therefore not useful for our purposes.

**Lacunarity**

Lacunarity is measure of translational and rotational invariance in the foreground shapes of an image. Assume \( r \) = box size, and \( S \) = box mass (built-up area pixels). We create a frequency distribution of box masses \( n(S,r) \) then convert to a probability distribution \( Q(S,r) \) by dividing each frequency value by the total number of gliding boxes of a given size \( N(r) \). We then calculate: \( Z(1) = \sum S Q(S,r) \) and \( Z(2) = \sum S^2 Q(S,r) \) (Filho & Sobreira, 2007). Lacunarity is thus a scale-dependent measure defined as:

\[
\Lambda(r) = \frac{Z(2)}{[Z(1)]^2},
\]

where higher heterogeneity is expected to produce higher values (Ibid, 4). Figure 2 is an example of how a single sub-scene of a formal settlement was transformed from its false-color composite through creation of the NDVI, creating a density slice of all values < 0.09 and then inverting the Black/White values to produce pseudo built-up (foreground) areas to compute lacunarity.

![Figure 2. Image transformation for lacunarity and fractal analysis.](image)

Table 1 lists the results for \( \Lambda \) and \( D_B \) for the settlements in the study area relying primarily on our estimate of pseudo built-up areas.
Table 1. Lacunarity and fractal related values by settlement type

<table>
<thead>
<tr>
<th></th>
<th>Foreground $\Lambda=1+c\nu^2$</th>
<th>Count of Boxes with FG Pixels</th>
<th>Fractal ($D_f$), Binary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Box Size 200, Binary</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Informal</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satellite</td>
<td>1.088</td>
<td>3187</td>
<td>1.877</td>
</tr>
<tr>
<td>Berlin</td>
<td>1.059</td>
<td>536</td>
<td>1.827</td>
</tr>
<tr>
<td>Joya</td>
<td>1.123</td>
<td>4792</td>
<td>1.806</td>
</tr>
<tr>
<td>LoDeCoy</td>
<td>1.059</td>
<td>295</td>
<td>1.665</td>
</tr>
<tr>
<td>Peronia1</td>
<td>1.044</td>
<td>2933</td>
<td>1.907</td>
</tr>
<tr>
<td>Peronia2</td>
<td>1.144</td>
<td>2157</td>
<td>1.833</td>
</tr>
<tr>
<td><strong>Formal</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Terrazas</td>
<td>1.045</td>
<td>2285</td>
<td>1.913</td>
</tr>
<tr>
<td>Balcones</td>
<td>1.022</td>
<td>3042</td>
<td>1.917</td>
</tr>
<tr>
<td>Mirador2</td>
<td>1.056</td>
<td>1178</td>
<td>1.94</td>
</tr>
<tr>
<td>GTowns2</td>
<td>1.047</td>
<td>592</td>
<td>1.89</td>
</tr>
<tr>
<td>GTowns3</td>
<td>1.042</td>
<td>510</td>
<td>1.806</td>
</tr>
<tr>
<td>Pinares</td>
<td>1.074</td>
<td>7856</td>
<td>1.908</td>
</tr>
</tbody>
</table>

As some of the prior literature showed, lacunarity values were generally higher in the informally settled areas, and based on the total number of sliding boxes for which $\Lambda$ was computed, there was a statistically significant difference between settlement types, indicating this measure could potentially be used to discriminate settlement types between higher or lower economic status, assuming residents in informal settlements exhibit lower socio-economic status than in planned communities. There was only an 87% chance the parent means (using the unpaired t-test) from the grayscale built-up images were from different underlying populations, while the binary image performed significantly better with $p(0.052)$, leading us to conclude that lacunarity calculations on binary image built-up areas exhibit settlement type discriminatory power, even in Guatemala. However, given the pre-processing required, the image conversion steps, the variability in selection of thresholds, box sizes, number of pixels to shift, and how to handle the edges in a masked image scene, this metric is somewhat arcane and not intuitive for widespread use.

Vegetation

The expected differences in settlement vegetation characteristics were patch size, percent vegetation, and compactness ratio. Patch size was expected to be smaller in the informal, economically disadvantaged settlements consistent with the research on greenspace as a contributor to quality of life. Only the vegetated areas inside the settlement boundary were measured, representing greenspace of the in-between spaces among the built-up areas. Total vegetation coverage was also expected to be less in the informal areas, assuming consistency with prior research on quality of life. It was assumed patch compactness would be greater, with patches having more circular shapes in the disadvantaged unplanned neighborhoods vs. elongated shapes from border-like patterns created by intentional and possibly ornamental planting in more affluent settlements. Figure 2 visualizes that more circular shapes exhibit greater compactness, as computed by:

$$\text{Compactness} = \frac{A_i}{B_i}$$

where $A_i = \text{Area of Patch}$, $B_i = \text{Area of Circle with same circumference as patch}$. Table 2 lists the vegetation-related results, allowing us to conclude that patch area, mean compactness ratio, and total % vegetation inside the settlement mask exhibit a statistically significant difference between settlement types at the $p(0.088)$, $p(0.00001)$, and $p(0.0099)$ levels, respectively. Despite the small apparent variation in the vegetation patch area and compactness ratio measures among formal and informal settlement types, the large number of patches from which these measures are derived still yield statistically significant differences in the samples.
Table 2. Vegetation related values by settlement type

<table>
<thead>
<tr>
<th>Settlement Type</th>
<th>Area in m²</th>
<th>Count</th>
<th>Compactness Ratio</th>
<th>Total % Veg</th>
</tr>
</thead>
<tbody>
<tr>
<td>InFormal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satellite</td>
<td>8.7</td>
<td>3151</td>
<td>0.73</td>
<td>6.6</td>
</tr>
<tr>
<td>Berlin</td>
<td>7.7</td>
<td>783</td>
<td>0.75</td>
<td>6.1</td>
</tr>
<tr>
<td>Joya</td>
<td>7.9</td>
<td>5503</td>
<td>0.74</td>
<td>7.1</td>
</tr>
<tr>
<td>LoDeCoy</td>
<td>10.9</td>
<td>480</td>
<td>0.74</td>
<td>10.1</td>
</tr>
<tr>
<td>Peronia1</td>
<td>8.8</td>
<td>2831</td>
<td>0.74</td>
<td>5.8</td>
</tr>
<tr>
<td>Peronia2</td>
<td>5.7</td>
<td>1783</td>
<td>0.75</td>
<td>3.9</td>
</tr>
<tr>
<td>Formal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Terrazas</td>
<td>8.4</td>
<td>2578</td>
<td>0.72</td>
<td>6.8</td>
</tr>
<tr>
<td>Balcones</td>
<td>10.4</td>
<td>5052</td>
<td>0.71</td>
<td>10.3</td>
</tr>
<tr>
<td>Mirador2</td>
<td>10.2</td>
<td>1885</td>
<td>0.72</td>
<td>10.7</td>
</tr>
<tr>
<td>GTowns2</td>
<td>8.4</td>
<td>1121</td>
<td>0.72</td>
<td>10.1</td>
</tr>
<tr>
<td>GTowns3</td>
<td>11.2</td>
<td>1021</td>
<td>0.72</td>
<td>12.9</td>
</tr>
<tr>
<td>Pinares</td>
<td>14.1</td>
<td>9226</td>
<td>0.71</td>
<td>12.1</td>
</tr>
</tbody>
</table>

CONCLUSION

This research has applied some measures found in the scientific literature that have been used to reveal whether there are differences in structure between rich and poor settlements, otherwise known as planned vs. informal typologies. It has merged information derived from high resolution imagery regarding built-up areas and vegetation to discern how some of the basic qualities of human settlements might have discriminatory power that sheds light on quality of life and economic wellbeing of the residents of an area. In this study area in Guatemala City, we have shown that higher lacunarity values in general are found in the informally settled areas, that fractal dimension using a pseudo-built-up representation is less promising, but that some interesting geospatial qualities of the vegetated patches within and among living spaces can provide clues to residents’ living patterns. Future research is aimed at building a multivariate model that incorporates other metrics of topography and transportation networks to determine which uncorrelated measures overall contribute the most to discriminating settlement type by integrating GIS and remote sensing image processing techniques.

ACKNOWLEDGEMENTS

A special thanks to D. Wong, D. White, G. Terrie, R. López and A. Gonzalez for their patient support and willingness to help over the past year and a half while I performed this research.

BIBLIOGRAPHY

Digital Globe, original ©Quickbird image, 23 March 09, Guatemala City, Guatemala.


