Remotely Sensed Population Counts of Informal Settlements in Cape Town, South Africa

GEOG 455 - Nadine Schuurman
April.3/17
Jeff Wang, Matthew Sutherland, Nicholas Kostachuk, George Dill-Jones

Introduction:

Cities in developing nations are having a difficult time dealing with the influx of people from rural immigration within their own country as well as refugee immigration from other nations. This influx is causing rapid growth of informal settlements, and is straining the cities' ability to provide essential services to their people. As much as 30% to 60% of the urban population of most large cities in developing countries are living in informal settlements (Veljanovski, 2012) and these cities lack the data necessary to apply monitoring systems for the purpose of planning and management. Population estimates for these areas are highly variable and highly debatable – for example in the Kibera neighborhood of Nairobi, Kenya, the population estimates range from 170 000 to 1 million residents (Veljanovski, 2012). Estimates done via field surveys are time consuming and expensive, and pose a potential safety risk to field workers (Earthzine, 2011). Remotely Sensed imagery is one possible information source to help cope with this problem.

Remotely Sensed imagery provides a visual representation of the landscape from above. Platforms for collecting this type of data generally consist of aircraft (planes, helicopters, or unmanned aerial vehicles) or satellites. Until recently, aircraft-based methods of image acquisition was the only way to get imagery with adequate resolution to delineate individual dwellings, which was time consuming and costly. Recent advances in satellite sensor technology have led to a proliferation of relatively cheap and accessible imagery. A lot of satellite imagery is even available free of charge, which could be of great use to developing nations where budgets are already strained.

The objectives of this paper are two-fold: first, we will do a scoping review on the literature pertaining to the use of remotely sensed imagery for doing population counts of informal settlements, and second we will develop a protocol and attempt to use that protocol to estimate the population of the informal settlements in Cape Town, South Africa. The paper will start with some background information on what a scoping review is, and then we will present the scoping review itself. After that, we will explain the protocol we have developed, and finally we will attempt to apply the protocol. The paper will conclude with a discussion of our application of the protocol, including any inadequacies or limitations found during the application process.

Background:

A scoping review is a type of literature review where the main goal is to relatively quickly map the key concepts of a research area. It also attempts to examine the types of evidence used and any gaps in the literature. While it is similar to a systematic review, there are some key differences. First, a scoping review takes a broad approach to reviewing the literature by trying to answer a relatively general question, while a systematic review tends to focus on a well-defined research question (Arksey, 2005). Furthermore, a scoping review attempts to represent all of the research from a specific field, whereas a systematic review will use only the best research to answer the question, and may largely ignore a large percentage of studies in favour of the ones that best fit the research question (Arksey, 2005). One particular aspect of scoping reviews draws a fair bit of criticism. When the framework was introduced, one of the defining characteristics of a scoping review was the fact that it does not attempt to assess the quality of research (Arksey, 2005). On one hand, this means they provide a relatively quick way to examine the full spectrum of knowledge and evidence in a specific field (Khalil, 2016) because assessing studies would add to the complexity and time commitments of such a review, while on the other hand it is argued that the lack of quality assessment means there is limited usefulness to policymakers (Levac, 2010). Daudt et al. (2013) argue that assessing the quality of research is a necessary step, and should be utilized to formulate the inclusion/exclusion criteria.

Arksey and O'Malley put fourth four reasons to undertake a scoping review [table 1]. We are undertaking this scoping review to summarize and disseminate research findings. Specifically, we are summarizing the research done to date about using remote sensing for population counts of informal settlements, and then using that to formulate a protocol that could be applied by others for the purposes of doing their own population estimations.

The first stage of a scoping review is to identify the research question (Arksey, 2005). This first must be done to guide the search strategies that will be developed in the next step (Arksey, 2005). Our research question is: "what has been written about using remotely sensed data to do population counts of informal settlements?" In our research question remote sensing data refers to imagery collected via satellites, UAVs, or aircrafts, and informal settlements are any settlements that are built without permits or permission and lack some or all of the typical services one would associate with formal settlements (such as running water and electricity.) This can include slums, shantytowns, and refugee camps as well as backyard residences.

The second stage of a scoping review is to identify studies that are relevant to the research question. A scoping review is supposed to be as comprehensive as possible, but practical issues relating to available time, funding, and access to resources often require researchers to find a balance between feasibility and breadth (Levac, 2010). It is during this step that inclusion/exclusion criteria and search strategies are developed (Daudt, 2013). There should be justified rationale included in the background of your review for the inclusion/exclusion criteria (Khalil, 2016). Our search strategy was to search three online databases (Web of Science, Environment Complete, Academic Search Complete) using a selection of search terms determined with help from our faculty Librarian [table 2]. As these are two databases specifically covering Geography topics, and one multi-disciplinary database we felt these should include a large amount of what is available on the study topic. The fact that there was a lot of repetition between databases reinforced our thinking. For our study, the inclusion criteria were that it was related to both remotely sensed data and informal settlements. The exclusion criteria were anything that was not available in full text because abstracts are not necessarily a full representation of an article, anything written before 1985 because the technology required would not have been available, and anything that was written in a language other than English because we did not have the resources necessary to deal with translation. We choose to include articles that are not only about population counts, but also about identifying informal settlements because it may be necessary to identify the locational extents of an informal settlement before attempting to determine the population within them.

The third stage, study selection, is where you apply your inclusion/exclusion criteria to the articles found during the previous step. This is important due to the need for the search strategy to be broad enough to include most if not all of the research in a field, which in turn produces a number of irrelevant studies (Arksey, 2005). These steps should be done simultaneously by at least two researchers to ensure rigor, and to reduce uncertainties in which studies tend to include (Levac, 2010). The first stage is for researchers to read the article titles to identify the papers eligibility (Archibald, 2016), such as removing foreign language titles and anything that is not related to both remote sensing and informal settlements, as well as duplicates. If relevance is unclear at any step, articles should be included in the next step of the screening process. Once article titles have been screened, abstracts are read and the inclusion/exclusion criteria are applied. If eligibility cannot be determined from the abstracts then the full article is read (Arksey, 2005). We had all four team members execute searches and select relevant studies based on title, and then each paper selected was reviewed by two team members. This narrowed our initial results of 453 to a final number of 27.

The fourth stage is to “chart” the data. Charting is a technique to sort material by key issues and themes (Arksey, 2005), and is useful for dealing with qualitative data. When deciding on what to chart, you must again balance comprehensiveness with feasibility. You need to extract the information necessary to answer the study question, without being too time consuming (Daudt, 2013). One problem that can arise at this stage is that it can be difficult to summarize complex concepts in a meaningful way (Levac, 2010). Conducting a trial exercise followed by a group consultation can provide the team with a valuable opportunity to reach consensus on what to chart (Daudt, 2013). Our team chose to chart author, date, study location, data used, method used, aim of study, and important results.

The final stage of the scoping review is to summarize and report the results. A systematic review will typically include reading a large number of studies but only report on the select few that best answer the research question, while a scoping review aims to produce an overview of all material reviewed (Arksey, 2005). Researchers must consider the best approach to state the results of a study, and how to best present the study findings to the reader (Levac, 2010). If working in a large team, it is advised to designate a smaller group to make meaning out of the data (Daudt, 2013). Our results are reported in table 3.

A consultation with stakeholders is a valuable tool to enhance a scoping review (Arksey, 2005), and should be considered a required component (Levac, 2010). To maximize the effectiveness of the consultation, the researchers should identify the purpose of the consultation beforehand. This is because without a clear purpose outlined first it may be unclear when, how, or why to consult with stakeholders (Levac, 2010). One option is to invite suitable stakeholders to be part of the research team, but not all stakeholders are appropriate to be involved as part of the research team (Daudt, 2013). For our study, we have been in constant contact with our professor and TA who are two of the major stakeholders in our research project.

Our study location is in Cape Town, South Africa. Cape Town is the largest city in the Province of Western Cape, and has an estimated population of 3.96 million people, up from the 3.7 million people reported in the 2011 census. An average household is comprised of 3.17 people. There are now almost 1.3 million households in Cape Town, with a substantial portion of new households being informal settlements (Integrated Annual Report, 2016).

Scoping Review:

Informal settlements are officially defined as those that exist in urban areas without formal approval or documentation. Such settlements can contain thousands of dwellings, such as large slum complexes, or may consist of individual buildings built on otherwise legally owned land (Ioannidis et al. 2009, Snyder et al. 2013.) Informal settlements are generally characterized by poor access to health services, limited access to education facilities, inadequate infrastructure, uncontrolled population densities, and they are often built in unsuitable and sometimes dangerous environments. Informal settlements are common in less developed countries where high populations and increasing demand for infrastructure cannot be matched with formal construction. In South Africa, around ten percent of the population live in informal settlements, while in Cape Town specifically, about twenty percent of households live in informal structures (Lehola, 2012). It is important to identify the people living in these informal areas so urban planning and resource distribution can be more effective. There are a number of methods, each with their own pro and cons, which can be used to estimate population in informal settlements.

Surveys

The simplest way to determine dwelling and population counts is through on the ground surveying of informal areas. In Rio de Janeiro, Brazil, urban informal settlements are often underrepresented in the national census. However in 2010 the Brazilian census sought to include informal settlement data more rigorously into its collection scheme. This survey found that about 22 percent of Rio's population (1.4 million people) lives in informal dwellings, and that as expected these areas are characterised by suboptimal living conditions (Snyder et al. 2013). Similarly, in 2000 the city of Pune in India requested the creation of a complete list of slum dwellers from the NGO Shelter Associates (SA). The resulting Pune Slum Census became the first socio-economic and spatial census to be created and analysed using GIS, providing the city with the basis for future development and urban planning (Joshi et al. 2003). If survey data for a given study area does not exist, one must either conduct one's own survey (requiring a whole host of alternative resources, volunteers, formal paperwork, etc.) or rely on alternative remote sensing methods in order to determine an informal settlement population approximation.

Manual Counts

Sometimes a study area is small enough to facilitate a manual count of dwellings and the determination of population based on previous survey data (Kakembo and van Niekerk, 2014, Chen et al. 2003). For the municipality of Nelson Mandela Bay in South Africa aerial photography from 1996 and 2006 was compared and analysed in order to derive a method for more accurate collection of census data. The spatial resolution was high enough and the study area was small enough to allow for the manual counting of dwellings (Kakembo and van Niekerk, 2014). For eleven sites in Bangladesh and several African countries, manual counts of residential structures were made from satellite imagery. Publicly available reports were used to estimate the mean number of people per structure, which were then applied to the counts in order to obtain an approximate population for the sites (Checchi et al. 2013). As part of a natural risk assessment of bush fires in Australia, dwellings were identified and manually digitized. This data was used to estimate the replacement cost of a house should it be destroyed in a fire event. As the study area for this analysis was relatively small (limited to areas prone to bushfire events), the manual identification of buildings was feasible (Chen et al. 2003). While the strategy of manual counts is very simple and could potentially be conducted by non-specialist volunteers, it is not practical for larger study areas and in such cases would be very time consuming and most likely unfeasible.

Population Growth Models

Population estimation methods are not limited by the availability of current data. It is also possible to use population growth models to estimate population using historical data. Al Mamun et al. (2013) conducted a land cover analysis on Dhaka City using Landsat Satellite images from 1990, 2003, and 2010, which applied supervised land use classification on each image (Al Mamun et al. 2013). Analysis of the images produced a rate of change for each individual land cover over time. It is possible to use this rate of change for land cover to forecast land cover changes using a cellular automata model (Al Mamun et al. 2013). It is then possible to use the land cover data for residential land use to estimate the population in a particular region. Olena et al. (2011) also applied a similar method using logistic regression modeling to predict the growth of informal settlements using data from 1990, 1995, 2000, and 2005 and was able to construct a model to determine population density for any given year in Istanbul, Turkey. The biggest issue in using population growth models is deciding which model is best for a particular purpose and study area. It is difficult to justify which model is best because "there is no unified approach to perform such analyses, as their objective, constraints, and assumptions are different" (Olena et al. 2011). In most cases, the "trial and error" (He et al. 2008) approach is used to find the best model for analysis for each individual study. One important thing to note is that informal settlements are usually associated with "rapid growth of urban population caused by rural immigration" (Velijanovski et al. 2012) and immigration happens to be one of the major limitations of population growth models.

Pixel Based Methods

Pixel based classification is another method by which the population in informal areas can be estimated. Pixel-based methods involve the assignment of individual pixels to different classes (vegetation, water, urban cover, etc.) based on their spectral signatures. The spectral signature for each pixel depends on the variation of its reflectance and emittance of electromagnetic energy. Both soft and hard pixel classification methods are possible. Hard pixel classification requires that each pixel to belong only to one land cover class, while soft pixel classification allows for the incorporation of fuzzy rules, with individual pixels potentially belonging to more than one class which is fairly common in remotely sensed data. While soft classification is a more realistic representation of land cover on the earth's surface, it does not identify specifically where the land covers are present within each pixel, and thus hard pixel classification is more useful for population estimation (Alpin and Atkinson, 2001).

Pixel-based methods have been used to examine settlement populations in a number of contexts, including formal (Stuckens et al. 2000, Alpin and Atkinson 2001) and informal situations (Engstrom et al. 2015, Bjorgo 2000). Of particular interest is the method of area sampling which has been used to rapidly assess population in disaster areas (Brown et al. 2001). After a region of interest is mapped or satellite imagery is acquired, the total area population is approximated by counting the number of people living in a limited number of blocks in the study area. The average population density ratio of these blocks is then applied to the entire area of the study site to get a population estimate for the whole region. Depending on the camp, the dimensions of the blocks varied from 25 to 100 meters squared. In situations where other data is not available or data quality is limited and other methods cannot be used, this area sampling method can be used to obtain a population estimate for an informal settlement.

Object-based image Analysis

Object-based image analysis is a common and useful analytical tool to estimate population with high quality remote sensing data. The relationship between the objects and spatial resolution can determine which image analysis is most appropriate. When the pixels are significantly bigger then the objects or the pixel and objects are close to the same size (low resolution), a pixel-based approach is more effective. If the object is significantly larger than the pixels, (high-resolution) object-based image analysis is the most appropriate (Blaschke, 2010). The rise in popularity of object-based image analysis can be contributed to an increase in high resolution data. With the launch of high resolution sensors on relatively new satellites such as Quickbird (1999) or IKONOS (2003) it has increased the accessibility of available high quality data. Another feature of the imagery provided by these satellites sensors are the ability to capture bands (Blaschke, 2010.) An emphasis on high quality data will be demonstrated throughout the analyzed studies. The following will review the process in an object-based image analysis.

Performing an object based image analysis has two main parts. The first part is image segmentation which divides the image into spatial segments. The goal is to have each one of these spatial segment to correspond to one real world object (Knudby, 2015). Objected- based image analysis is a technique where objects are segmented based on similar characteristics such as size or shape. This image segmentation has four main approaches: region based, edge based, point based or using a combination of these techniques (Blaschke, 2010). An object-based image analysis can be performed as a supervised or unsupervised classification. In an unsupervised classification, the software will break the image into segments based off of computer algorithms. This means the user does not get to determine the number of classes and needs to be able to determine what the computer made classes are. A supervised classification involves the user setting up “training sites”, which are areas in the image that are identified by the user. These “training sites” are then used by the computer to segment the rest of the image (Knudby, 2015). The following section will look at multiple studies using object-based image analysis to investigate informal settlements.

Population counts of informal settlements using object-based image analysis has been performed in various parts of the world. One area is the Kibera neighborhood in Nairobi, Kenya. In this study, object-based image analysis was used for two main goals. This first goal of the study was to derive a detailed land cover change map that is able to supply population estimates. Another goal of the study was to analyze the growth and changes between 2006 and 2009. Object- based image analysis was appropriate for this study because of the following reasons: First, high resolution data was available from GeoEye and QuickBird. The data from 2006 to 2008 was obtained from the GeoEye sensor which had 0.6 meter resolution. The data for 2009 was obtained from the Quickbird sensor and had a resolution of 0.5 meters. Both of the satellite sensors provided imagery with blue, green and red bands. Second, is that they need to be able to identify the difference between similar characteristic objects such as rooftops, unpaved roads and non-built land. This was done by supervised classification using ENVI and the Feature Extraction module. Different values for the segmentation and merging patterns were set to influence the classification results. The objects extracted during the segmentation were classified using Support Vector Machine (SVM) classification algorithm as well as expert advice. This study was conducted in a large area that needed to be broken into 12 parts due to data size. The results from this objected-based image analysis led to analyzing change detection and was not used for initial population counts. (Veljanovski et al. 2012)

Object- based image analysis was performed in another study in Rio De Janerio, Brazil. This study was attempting to estimate population in a squatter settlement using both 3D and 2D data derived from a digital elevation model. Although this study used 3D modelling approaches in conjunction with object-based image analysis we will only focus on the object-based image analysis aspect. Object-based image analysis was used to employ modelling strategies that include thematic or specific knowledge. In this study, the user inserted the high resolution data obtained from IKONOS satellite to discriminate nonresidential building from residential buildings. Another important aspect of this imagery is the red, blue, green band included. This is an advantage of object-based classification as it excels in segmenting spatially similar objects. This is useful when attempting to separate structures in informal settlements as many shacks could potentially look the same. Once the data was obtained from the object-based image analysis the population estimates were derived from population density data from the data supplied from the Pereira Passos Institute (IPP) in Rio (Almeida at el. 2011). After the object-based image analysis has been completed for Cape Town, South Africa, local public data will be consulted to get persons per household. If this data is not available, other resources will be investigated such as expert or other literature. As demonstrated by these studies, object-based image analysis can be an effective method for performing population counts within informal settlements. There are various way that object-based image analysis can be performed. The following will review various GIS and remote sensing software available to us to perform an object-based image analysis.

There are numerous software programs that have the potential to perform an object-based image analysis. A study performed in Accra, Ghanna used multiple GIS programs to perform object-based image analysis and compared the two methods. For this study, high-resolution data was obtained from the Quickbird satellite, which collects panchromatic imagery at 0.6 m and multispectral imagery at 2.4 m resolution. This imagery includes the blue, red and green bands. The purpose of this study was to determine the best method for land cover change. Although, we are attempting to perform population counts this research paper was still able to provide crucial insight. It compared two methods for determining land class change with object-based image analysis: post- classification comparison and bi- temporal layerstack methods. (Tsai et al. 2011) Both of these methods require multiple dates of remotely sensed data to produce results for land cover change. This study can be used to analyze the two different potential programs available to us to perform the population count: feature analyst in ARCGIS and feature extraction in ENVI. The results of the study determined that feature analyst in ArcGis is more effective due to its reliability. Feature analyst did a superior job in delineation of individual buildings of various sizes. This is critical information for our project as informal settlements often have random sized homes as well a little structural organization. A major concern in the study was identifying the difference between the different roof materials. This is relevant because to effectively perform a population count we will need to identify and count the number of households. Though this was an issue for both programs, feature analyst in ArcGis achieved a higher accuracy then feature extraction in ENVI. (Tsai et al. 2011) Overall, feature analyst did a better job using the “training sites” set up by the researchers to segment the rest of the image. In conclusion this study has provided important insight in determining which program to use.

Data

Initial attempts were made to acquire aerial photo data from Dr. Prestige Tatenda Makanga with 6cm by 6cm resolution. Due to confidentiality reasons, the data could not be released to us. An OpenStreetMap TIF file was instead used to attempt to count populations in Cape Town. Since OpenStreetMap doesn’t come with a projection system, when the image was imported into ArcGIS, additional steps were required to incorporate a projection system into the TIF file. The TIF file was georeferenced against the base map already present within ArcGIS. Twenty-five ground points were selected randomly to georeference the study area to the base map. The georeferenced TIF file was then clipped to the study area polygon given by Dr. Prestige Tatenda Makanga, which already identified areas that are informal settlements within Cape Town. Additionally, point data for all the houses within Cape Town was also given to provide us with reference data to evaluate our protocols.

Protocol

Our research indicated that object-based image analysis is most appropriate when the data is of sufficient quality. The data we obtained from Open Street Map had a spatial resolution of 0.6 meters but lacked an adequate number of bands to perform the correct analyst on the imagery. This protocol refers to the Feature Extraction tool in ENVI despite Feature Analyst in ArcMap being slightly more effective (Blaschke, 2010). This is due to the lack of current knowledge and the inability to obtain knowledge about Feature Analysts because of licensing limitations. Using Feature Analyst would allow us to experiment and figure out the best potential protocol. We believe that the protocols developed for ENVI's Feature Extraction will provide an overview of the steps that can be implemented in ArcMap's Feature Analyst. The suggested protocol is as follows:

Open high quality data in ENVI and run the Feature Extraction module. The ideal resolution depends on the object you are attempting to identify. Based off reviewed literature we suggest using imagery with the highest spatial resolution imagery possible with a maximum resolution of 2.5 meters (Veljanovski et al. 2012, Almeida at el, 2011.) Another requirement for the imagery is to have at least three bands. All of the imagery in the reviewed literature was taken from sensors that are capable of taking multispectral imagery. Two extraction methods can be used:

Ruled Based Extraction:

1) Select the "Rule Based Feature Extraction" tool from the feature extraction toolbox.
2) Select the study area image as the Raster File in the Data Selection window.
3) ENVI will then ask about several parameters that need to be set up:

a) The "Algorithm" needs to be selected. A choice of region based, edge based or point based segmentation is required. We suggest using edge based segmentation.

b) The segmentation can be previewed. The "scale level" is a preview of the computer set segmentation. By clicking preview you are able to check how the computer will segment the image when the tool has been run. This should be done throughout the process to ensure quality of the segmentation.

c) “Merge” will be used within the tool. The merging tool will correct errors by merging the smaller segments to the larger more appropriate segments. The merge algorithm should be set to “Full Lambda Schedule”.

d) The “Texture Kernel Size” should be set to 3.

4) When attempting to correctly segment the rooftops it is necessary to set up rules for the module so that is it able to properly segment the sections. These rules will generate the probability of class membership of each segments. We suggest the following rules:

Rule 1: Create a rule based off of spectral mean. Select the type to “spectral” and the name to “spectral mean” and “normalized differences”. Rooftops often lack vegetation therefore is it possible to segment based off of NDVI values. The values can be adjusted by dragging the green and blue lines (minimum and maximum values) on the graph.

Rule 2: Create a rule based off the shape of the houses. Select the type to “spatial” and the name to “rectangular fit”. Once again we can adjusted the values of how rectangular you want the segments with the blue and green lines.

Rule 3: Create a rule based off of size of houses. Select the type to “spatial” and the name to “area”. Experiment with the green and blue lines to get the area of the segmentation based off size matches those of the households.

Rule 4: Add a rule to remove the road segments. The type will be “spatial” and the name is “elongation”. Once again you will need to experiment with the minimum and maximum values to ensure you are getting rid of roads without getting rid of any of the buildings.

5) Finish the segmentation by clicking “finish” to the save the output. Based on the results, fine-tuning may need to be done which can be done by adjusting the thresholds (green and blue lines) or modifying the rules.

Example Based Extraction:

1) Import study area data into ENVI using the Open function.

2) Select “Example Based Feature Extraction” from the Feature Extraction toolbox.

3) Selected the study area image as the Raster File in the Data Selection window.

4) On the Object Creation window, used the preview window in order to determine the best setting for image segmentation. ENVI will ask to set up parameters, which will involve adjusting the algorithm types and sliders in order to find the most appropriate segmentation threshold (see above).

5) In the next window, choose examples for the classes to be created. We suggest setting up five classes which will be: dwellings, roads, water, vegetation and dirt covered areas.

6) In the final window, select the save locations for the classification outputs (shapefile and raster). Select “finish” to save the outputs.

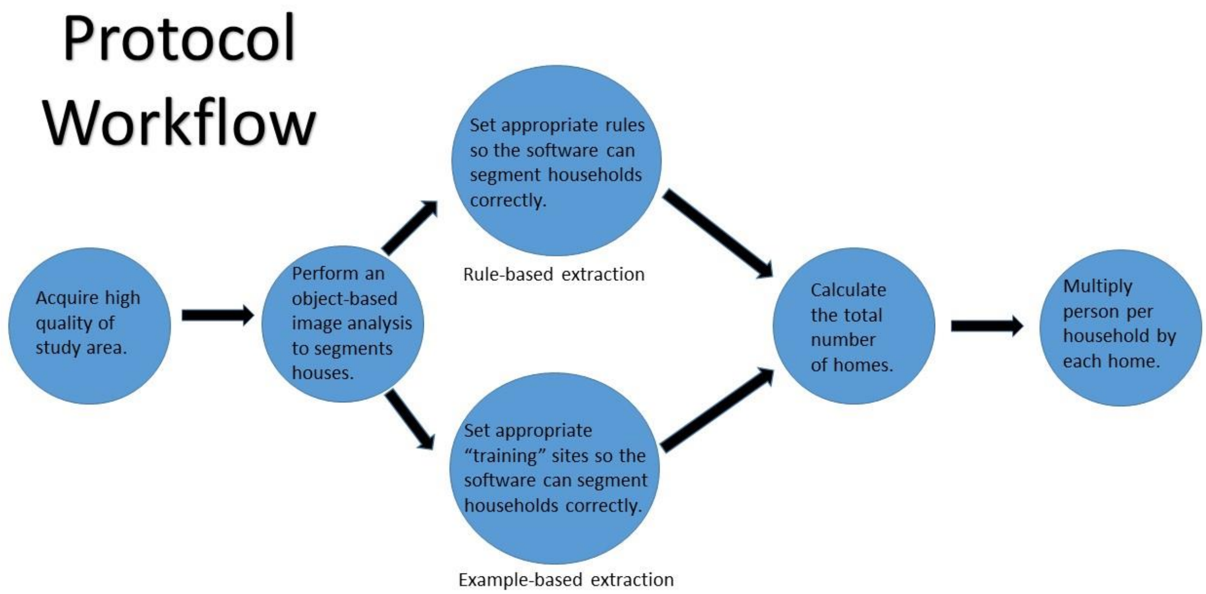

Figure 2. Ideal basic workflow to count populations in informal settlements.

Method

In our analysis, the area sampling by manual count method was used instead of an object-based approach due to data quality. The first step is to start with random samples of thirty rectangular polygons with an area of $100m^2$ to $800m^2$ in size. The thirty random sample sites are shown in Figure 1. The literature revealed that the adequate number and size of square blocks are still in question (Brown et al. 2001). By statistical principle, the representativeness of the blocks are proportional to the increase of blocks and the increase of the block size. It should be considered that as the number of blocks or the size of the blocks increase there will be diminishing return in the accuracy of the results. Thirty blocks was an arbitrary number chosen base on factors such as manpower, time constraint, and quality of data. The counting of population within the blocks can be done by interviewing all heads within the household through on site surveying, but such a method cannot be done due to physical limitations (Brown et al. 2001). We have to resort to counting houses manually and using the average persons per household to estimate population.

The informal settlements within Cape Town also have a varying degree of density and could possibly be taken into account by “observation while walking around the camp” (Brown et al. 2001). Since on site investigation is impossible for us, the assumption has to be made that the average density of all the polygons will be sufficient to cancel out the variations in house density. The house count is entered manually into the attribute table in ArcGIS and the area of the block is calculated automatically in m^2 with the geometry calculator. (Total area within blocks/total house count) is the average house density of the study area. The ratio is then extrapolated to the total area of the study area to get the estimated house count for the settlements. The estimated house count is used to compare to the point house data that is given to evaluate the effectiveness of the protocol. Finally, the estimated house count can be multiplied by the persons/house ratio found on Cape Town government website to find the estimated population for informal settlement in Cape Town.

Sample Sites in Cape Town

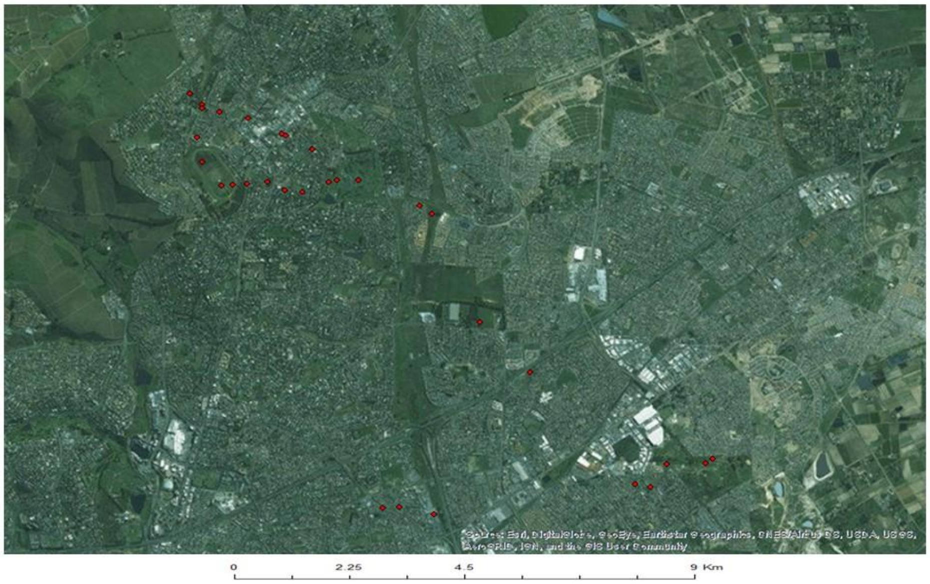

Figure 1. This imagery shows the random sample sites (shown in red) selected in pixel based approach in Cape Town, South Africa

Results

The area sampling by manual count approach produced a total area within the polygons of 13019.7m^2 and the total house count within the polygons was 155 houses. The ratio of house/area is 0.12 house/m^2 and the total area of study is was 3970808.4m^2. By extrapolating the ratio to the entire Cape Town study area, the protocol gives an estimation of 47273 houses. The reference point data for the houses shows a count of 43260 in total, which means that the relative error for the area sampling method is approximately 10 percent. The ratio of dwellers per house is 3.17, which means that the population estimate of the informal settlements in the study area is 149855 people.

Discussion /Limitations

The low quality of the data that was available played a huge role in the effectiveness of the protocol. Ideal data should contain multiple bands with at least 2.5m or smaller spatial resolution. This is especially true for our object-based protocol, as the lack of bands made the identification of houses extremely difficult. Both rule based and example based object-based image analysis attempted within ENVI were unable to distinguish houses from other objects in the study area. While it is possible to obtain other data, all that is available is lacking in either spectral or spatial resolution. A formal request to companies that sell higher quality data was also considered, but due to time and monetary constraints acquisition was not possible for this project. However, the object-based protocol should still be in theory the most effective method for counting population within informal settlements assuming that ideal data is acquired. The use of other programs to make up for non-ideal data were also considered. Feature Analyst from ArcGIS has been determined to have superior ability to distinguish objects than ENVI (Tsai et all, 2011). Unfortunately, the version ArcMap that we had access to did not have the licensing associated with Feature Analyst.

Area sampling by manual count was the second method by which we attempted to estimate the informal population within the study area. Even though the method is relatively simple when compared to object-based analysis, there are still a couple of things that can still be improved to further increase its accuracy. Manual counts are more effective with better spectral and spatial resolution, as the likelihood of miscounting houses is reduced. For our analysis only thirty sampling areas were chosen due to limited time and manpower. In statistics, thirty is considered to be statistically powerful (Brown et al. 2001). However, it would be possible to outsource time and labour to the general public to produce an even more powerful sampling. The act of manually counting houses can be done by anyone even if they do not have specialist GIS knowledge.

Conclusion

The protocols outlined in this paper are not designed to be used independently without the application of other techniques. The object-based protocol utilizes algorithms to distinguish houses within the aerial photography, but determining the number of houses within the study area is not sufficient to estimate the total population. Survey data taken from the Cape Town website on the average persons per house ratio has to be used in conjunction with object-based analysis to obtain the total population within the Cape Town informal settlements. Similarly, the area sampling by manual count protocol requires that each house to be counted manually within the blocks before it can be extrapolated to the whole study area. Again, the census data from Cape Town website for average persons per house ratio was used to obtain the total informal population. It should be noted that population growth models were not included in the protocols because of the lack of historical Cape Town data, as well as the lack of manpower to perform trial and error on different available models. Population growth models also fail to consider immigration, which is one of the major reasons for the growth of informal settlements. Pixel analysis is also not considered, as the primary strength for pixel analysis is to compute land cover and land use and there are no existing papers that use this method to calculate population estimates.

Ideally, the provided data should have a high enough spatial and spectral resolution to use an object-based protocol to estimate population. In our case, where object-based methods cannot be used effectively, it is possible to use an area sampling by manual count method instead. Since object-based methods were unable to give us any meaningful results, we are unable to compare our two protocols and examine any potential tradeoffs in accuracy. It is important to note that area sampling by manual count can be time consuming when applied to a larger area, but this limitation can potentially be solved by outsourcing the work to members of the public. With this in mind, future studies should attempt to evaluate the tradeoffs between object-based methods and area sampling by manual count when estimating the populations of informal settlements.

Search Terms
(slum* OR "informal settl*" OR shanty* OR "refugee camp*" OR "displaced persons" OR "displaced popul*")
("population count*" OR census OR population)
(GIS OR spatial OR map* OR aerial OR satell* OR "remote sens*")

Table 1: Search terms used

Reasons for undertaking a Scoping Review	Additional information
"to examine the extent, range and nature of research activity"	Map a field of study where the range of material available is difficult to visualize.
"to determine the value of undertaking a full systematic review"	Determine if a systematic review is feasible and/or relevant, and associated costs.
"to summarize and disseminate research findings"	Inform policy-makers, practitioners, consumers who lack time or resources to do a review.
"to identify research gaps in the existing literature"	Focus shifts from what has been covered in a field to what has not.

Table 2: Reasons for doing a scoping review, according to Arksey and O'Malley (2005).

Study Location	# of studies
Africa	12
Asia	9
South America	4
Europe	2
Australia	1
Data Used	
Satellite Imagery	19
Land Cover Data	2
Aerial Photography	4
census	4
SAR Imagery	1
Method Used	
Object-based analysis	9

pixel-based analysis	4
Survey into GIS boundaries	3
Manual counts	3
Total Studies: 27	

Table 3: Summary table for scoping review

Works Cited

Anders, Knudby. “Lecture 5: Object-based based image analysis” *Remote Sensing.* Powerpoint Slides. Simon Fraser University (Spring 2015)

Anders, Knudby “ Lecture 4: Classification” Remote Sensing. Powerpoint Slides. Simon Fraser University (Spring 2015)

Almeida, Cláudia Cléber G. Oliveira, Camilo D. Rennó, Raul Q. Feitosa “Population Estimates in Informal Settlements Using Object-Based Image Analysis and 3D Modeling.” Earthzine, August 16, 2011. https://earthzine.org/2011/08/16/population-estimates-in-informal-settlements-using-object-based-image-analysis-and-3d-modeling/.

Al Mamun, Abdullah, Asif Mahmood, and Mafizur Rahman. “Identification and Monitoring the Change of Land Use Pattern Using Remote Sensing and GIS: A Case Study of Dhaka City.” *IOSR Journal of Mechanical and Civil Engineering* 6, no. 2 (2013): 20–28.

Aplin, P., and P. M. Atkinson. “Sub-Pixel Land Cover Mapping for per-Field Classification.” *International Journal of Remote Sensing* 22, no. 14 (January 1, 2001): 2853–58.

Archibald, Daryll, Rebecca Patterson, Erna Haraldsdottir, Mark Hazelwood, Shirley Fife, and Scott A. Murray. “Mapping the Progress and Impacts of Public Health Approaches to Palliative Care: A Scoping Review Protocol.” BMJ Open 6, no. 7 (July 1, 2016)

Arksey, Hilary, and Lisa O’Malley. “Scoping Studies: Towards a Methodological Framework.” International Journal of Social Research Methodology 8, no. 1 (February 1, 2005): 19–32.

Bjorgo, Einar. “Using Very High Spatial Resolution Multispectral Satellite Sensor Imagery to Monitor Refugee Camps.” *International Journal of Remote Sensing* 21, no. 3 (February 15, 2000): 611–16.

Brown, Vincent, Guy Jacquier, Denis Coulombier, Serge Balandine, François Belanger, and Dominique Legros. “Rapid Assessment of Population Size by Area Sampling in Disaster Situations.” *Disasters* 25, no. 2 (June 2001).

Blaschke, T. “Object Based Image Analysis for Remote Sensing.” *ISPRS Journal of Photogrammetry and Remote Sensing* 65, no. 1 (January 2010): 2–16. doi:10.1016/j.isprsjprs.2009.06.004

Checchi, Francesco, Barclay T. Stewart, Jennifer J. Palmer, and Chris Grundy. “Validity and Feasibility of a Satellite Imagery-Based Method for Rapid Estimation of Displaced Populations.” *International Journal of Health Geographics* 12, no. 1 (January 2013): 1–12.

Chen, Keping, Russell Blong, and Carol Jacobson. “Towards an Integrated Approach to Natural Hazards Risk Assessment Using GIS: With Reference to Bushfires.” *Environmental Management* 31, no. 4 (April 1, 2003): 0546–60.

Daudt, Helena M. L., Catherine van Mossel, and Samantha J. Scott. 2013. "Enhancing the Scoping Study Methodology: A Large, Inter-Professional Team's Experience with Arksey and O'Malley's Framework." BMC Medical Research Methodology 13 (1): 1–9.

Dubovyk, Olena, Richard Sliuzas, and Johannes Flacke. "Spatio-Temporal Modelling of Informal Settlement Development in Sancaktepe District, Istanbul, Turkey." *Isprs Journal of Photogrammetry and Remote Sensing* 66, no. 2 (March 2011): 235–46.

Engstrom, Ryan, Avery Sandborn, Qin Yu, Jason Burgdorfer, Douglas Stow, John Weeks, and Jordan Graesser. "Mapping Slums Using Spatial Features in Accra, Ghana." *2015 Joint Urban Remote Sensing Event (Jurse)*, 2015.

Ioannidis, Charalabos, Christodoulos Psaltis, and Chryssy Potsiou. "Towards a Strategy for Control of Suburban Informal Buildings through Automatic Change Detection." *Computers, Environment and Urban Systems* 33, no. 1 (January 2009): 64–74.

Joshi, Pratima, Srinanda Sen, and Jane Hobson. "Experiences with Surveying and Mapping Pune and Sangli Slums on a Geographical Information System (GIS)." *Environment and Urbanization* 14, no. 2 (October 1, 2002): 225–40.

Kakembo, V., and S. van Niekerk. "The Integration of GIS into Demographic Surveying of Informal Settlements: The Case of Nelson Mandela Bay Municipality, South Africa." *Habitat International* 44 (October 2014): 451–60.

Khalil, Hanan, Micah Peters, Christina M. Godfrey, Patricia McInerney, Cassia Baldini Soares, and Deborah Parker. "An Evidence-Based Approach to Scoping Reviews." Worldviews on Evidence-Based Nursing 13, no. 2 (April 1, 2016): 118–23.

Levac, Danielle, Heather Colquhoun, and Kelly K. O'Brien. 2010. "Scoping Studies: Advancing the Methodology." Implementation Science 5: 69.

Snyder, Robert E., Guillermo Jaimes, Lee W. Riley, Eduardo Faerstein, and Jason Corburn. "A Comparison of Social and Spatial Determinants of Health Between Formal and Informal Settlements in a Large Metropolitan Setting in Brazil." *Journal of Urban Health* 91, no. 3 (June 1, 2014): 432–45.

Stuckens, J., P. R. Coppin, and M. E. Bauer. "Integrating Contextual Information with per-Pixel Classification for Improved Land Cover Classification." *Remote Sensing of Environment* 71, no. 3 (March 2000): 282–96.

Tsai, Yu Hsin, Douglas Stow, and John Weeks. "Comparison of Object-Based Image Analysis Approaches to Mapping New Buildings in Accra, Ghana Using Multi-Temporal QuickBird

Satellite Imagery." International Journal of Remote Sensing 3, no. 12 (December 16, 2011): 2707–26.

Veljanovski, Tatjana, Urša Kanjir, Peter Pehani, Krištof Oštir, and Primož Kovačič. "Object-Based Image Analysis of VHR Satellite Imagery for Population Estimation in Informal Settlement Kibera-Nairobi, Kenya," 2012.

"Population Estimates in Informal Settlements Using Object-Based Image Analysis and 3D Modeling." Earthzine, August 16, 2011. https://earthzine.org/2011/08/16/population-estimates-in-informal-settlements-using-object-based-image-analysis-and-3d-modeling/.

"City of Cape Town Integrated Annual Report, 2015/2016." City of Cape Town. http://resource.capetown.gov.za/documentcentre/Documents/City%20research%20reports%20and%20review/City%20of%20Cape%20Town%202015_16%20Annual%20Report%20Final_A.pdf