Waffle Homes: Utilizing Aerial Imagery of Unfinished Buildings to Determine Average Room Size

Carson Woody
Human Geography Group, Oak Ridge National Laboratory, TN, USA

Tyler Frazier
Human Geography Group, Oak Ridge National Laboratory, TN, USA

Abstract

A primary function of the Population Density Tables Project (PDT) at Oak Ridge National Laboratory is to produce residential population densities per 1000 sq. ft. for each country and their associated first-level administrative units. This is accomplished by utilizing the average size of different types of dwelling areas (urban, rural, single-family, multi-family, etc.) and the average household size provided by a country's Census or statistical bureau records. This data is available for the majority of Europe, North America, and large swathes of Asia, but is less easily found in Africa and South America. In these regions, Censuses generally report dwelling area by number of rooms, which poses the challenging question of how we can translate number of rooms to dwelling size when no dwelling size areas are available with which to compare. Using sub-meter resolution satellite imagery of Accra, Ghana, this challenge can be tackled using imagery of roofless buildings currently under construction that show the interior floor plan of the dwelling. A sample of buildings from the different neighborhoods of Accra can be digitized to provide an estimate and range of average room sizes of dwellings. This average room size can then be translated to a total dwelling area using the "number of rooms occupied by a household" variable from the Ghanaian Census. This intermediate step between average dwelling size and number of rooms occupied, fills the missing link that prevents PDT from continually producing new population densities for countries where dwelling size is unavailable through any official means.

2012 ACM Subject Classification Social and professional topics → Cultural characteristics

Keywords and phrases Urban Analytics, Aerial Imagery, Satellite Imagery, Population Density, Human Geography, Africa, Residential Dwellings

Digital Object Identifier 10.4230/LIPIcs.GIScience.2023.85

Category Short Paper

Acknowledgements This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (`http://energy.gov/downloads/doe-public-access-plan`).

1 Introduction

Oak Ridge National Laboratory's Population Density Tables (PDT), is an information system with a graphical user interface that measures population density for over 60 facility types by people per 1000 sq. ft. [9]. Generally, this is performed under a Bayesian approach using

Figure 1 Waffle Home Example (Accra, Ghana) [4].

prior knowledge from expert elicitations and from gathering new data on individual facility types in a region to update the model [10]. Uncertainty is propagated throughout each model, starting at the data entry stage, where a range of data can be input into the PDT interface. Using this approach, PDT estimates ambient building occupancy at the national and subnational level. To model residential building occupancy, PDT uses the RevengC R package to reverse engineer census data and produce an uncensored table of population density probabilities. This R package is used in PDT in the form of the Census Tool, which uses average household size and average dwelling size as inputs to produce a residential dwelling density per 1000 sq. ft. These inputs are primarily obtained from country statistical bureau produced censuses or statistical documentation as PDT views Census data as the most authoritative sampling source [2].

Dwelling sizes are generally accounted for in national censuses and state statistical bureaus in the form of floor area in sqm. or sq. ft. However, in the majority of Africa and South America, recent censuses account for dwelling size in the form of "number of rooms occupied" in a dwelling unit. Without knowing the actual size of these rooms occupied, it poses a challenge to measure residential population density for a large portion of the world.

To use the data provided from censuses where dwellings are measured by number of rooms occupied as opposed to floor area, a link between the rooms inhabited and the actual size of these rooms is needed. We found this intermediary in Accra, Ghana (AOI), and its large volume of unfinished "honeycomb" structures where one can easily see the interior layout of the building as seen in Figure 1. These were coined "waffle homes" due to their grid iron like appearance, and along with Accra, were found in several other large African cities, like Lagos, Khartoum, and Conakry. From this type of imagery, one can easily identify the separate rooms in a building. Subsequently, individual rooms can be digitized using GIS software. With a statistically robust sample of digitizations, an average room size can be defined.

2 Related Works

Traditionally, the compound house is one of the most common typologies of housing in Accra, and tends to be the subject of research in this area. This is a common home style for lower income households and generally consist of a series of single hall units surrounding an open courtyard with shared kitchen and bathroom facilities, and generally cover an area of 100 sqm. [1]

However, compound homes have been losing popularity and are considered outdated as households prefer single-family living structures, which encourages builders to prioritize more single-family homes and apartments. This trend coupled with the continued influx of people to Ghana's urban areas, has increased the need for housing and rooms built, with an estimated 5.7 million rooms having been required to house the population by 2020 [6]. Many compound homes are not expected to survive the new wave of construction as new and affordable housing targets the middle class [3].

This wave of new building can be seen in the aerial imagery of Accra. Little research could be found using the aerial imagery of under construction homes to determine dwelling and average room sizing, but there has been research using architectural floor plans to study the general area of different types of rooms in different housing layouts [5]. This work relied on published floor plans from New Zealand to measure room area and went further by differentiating rooms by function and floor layout. Our study only relies on the aerial imagery of under construction dwellings without the ability to ascertain each room's individual function, but the New Zealand study did show that knowing the average area of different rooms of a home can lead to an accurate estimate of the home's total floor area [5]. While building typology differs greatly between New Zealand and Accra, the same idea of measuring room size can be used to find dwelling areas in our AOI.

3 Methodology

The Greater Accra Region, a subnational unit of Ghana including the country's capital, was used as the AOI for this new methodology. This region works as an ideal case study due to its lack of a census designated average dwelling floor area and its large number of waffle home type structures from which to sample. In addition, the 2021 Ghana Population and Housing Census recently became available which provided the necessary data on dwelling size by number of occupied rooms as well as average household sizes. This number of rooms occupied data provided by the 2021 Ghana Census will be used in conjunction with the waffle home digitizations to determine an average floor area [7].

Samples of waffle homes were gathered using Digital Globe/Maxar imagery from the Greater Accra Region. Samples were manually identified using imagery tiles from each district of the Greater Accra Region, and a total of 1267 sample points were identified by their unique grid-iron appearance in aerial imagery. A point data set was created from these structures, and the embedded PDT smart sampling tool was used to create a statistically robust selection of buildings across Accra. The embedded smart sampler tool works by randomly selecting points from the data set to digitize until the sample is large enough to closely represent the "true average" of the initial data set. It does this by having the data set pass three statistical tests before being considered a statistically robust selection of the data set. This was done due to time constraints and to lessen the number of buildings and rooms that needed to be digitized to create a data set of room sizes.

The first test the embedded smart sampler tool uses is a consistent distribution check, which ensures the data has the same estimated distribution, in this case a log-normal distribution. Then, the data set must pass a mean percentage change check, to ensure that as each new data point is sampled, the average is moving closer to the true average. An acceptable mean percentage change was set at 5 percent for each time a new data entry is added, and once the entries fall within this threshold a specified number of times, it passes this test. The last test, a Margin of Error check, sets a maximum margin of error, here it is 10 percent, and the embedded tool will calculate a range of values using a confidence interval of 95 percent. Once each new entry falls within the margin of error range a specified number of times, it will pass test three.

A data set of 697 rooms was created using the smart sampler tool by randomly selecting from the 1267 waffle house points to digitize the areas of each defined room in the dwelling. The 2021 Ghana Census's definition of an "occupied room" was used, which states that occupied rooms include living rooms, bedrooms, sleeping rooms, and dinning rooms, but excludes closets and bathrooms. This definition was used along with the 2018 Ghana Building Codes stipulation that an occupied room shall not be less than 9.5 sqm. to uniformly remove any bathrooms/closets from the room data set.

4 Results

This new data set of 697 rooms averages out to be 20.38 sqm. per room. Throughout this process, an additional 45 "waffle house" points were sampled and each of the rooms were digitized during further testing and research. This brought the total rooms data set up to 1089 rooms, and a new room average of 20.37 sqm. per room. Even with the addition of 393 rooms to the data set, the change in area was only .01 sqm. This supports the assumption that expanding the survey to additional points across the region would still return a similar average room size.

The resulting sample data set of waffle-home rooms shows a log-normal distribution with the majority of samples falling under 30 sqm., reaching a cumulative frequency of 85.49 percent at 30 sqm. The majority of the samples themselves fell between 10 to 20 sqm with 578 of the 1089 sample areas. Figure 2 shows the area distribution of the data set.

The Ghana Living Standards Survey of 2008 (GLSS5) was used as a comparison metric. While dated, it provides the average dwelling floor area for the Greater Accra area that the 2021 Ghana Population and Housing Census lacks [8]. The Ghana Population and Housing survey was used to determine the average area occupied by using the average number of rooms occupied (2.09) of the Greater Accra Region (GAMA), along with the newly determined average room size [7]. Table 1 shows a comparison of average areas between the two studies.

Table 1 GLSS5 and Waffle Home Data set Comparison.

	GLSS5 Urban [8]	GLSS5 GAMA [8]	Waffle Homes GAMA
Average Room Size	19.59 sqm	23.66 sqm	20.37 sqm
Average Area Occupied	33.3 sqm	42.6 sqm	42.57 sqm

The results produced by the Waffle Home methodology are similar to those produced by the GLSS5. Validation is difficult in this scenario as there are no recent state produced dwelling area averages against which to compare for Ghana or from the neighboring countries. More data will be needed, either from a state sponsored survey or additional waffle home POIs, to provide a more rigorous validation.

PDT's future goal is to build upon our current Bayesian modeling process to utilize this new method of sampling room floor areas along with the "rooms occupied" tables provided by state-level censuses to create total dwelling areas. PDT observation models capture

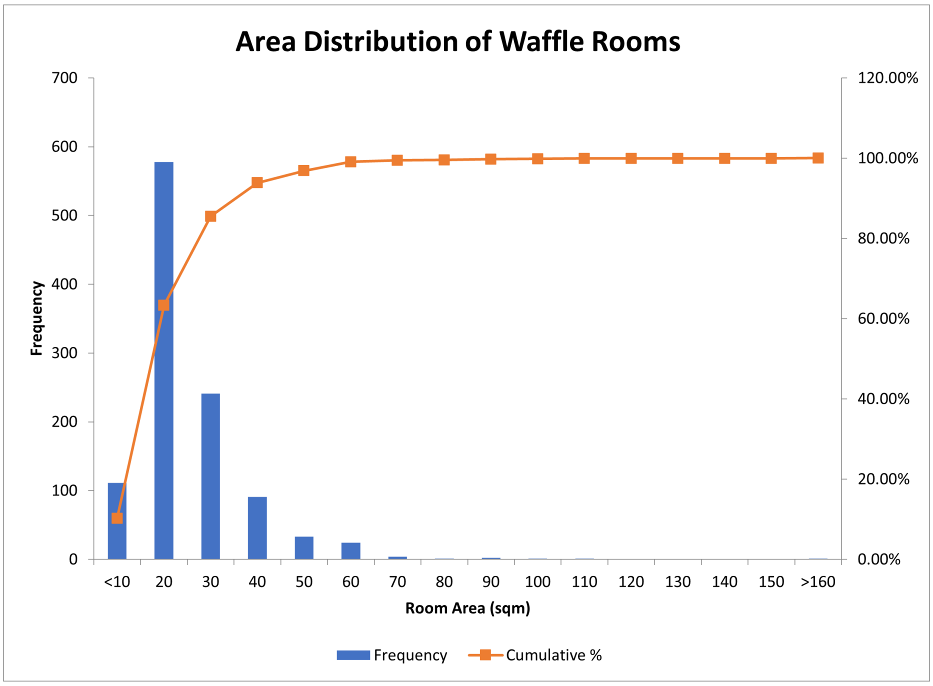

Figure 2 Histogram and Cumulative Frequency of Waffle Home Data set.

uncertainty at the data entry level and propagate it through the model so it can be observed in the resulting probability density function. This feature will be necessary in the model expansion for this format of data to capture the uncertainty associated with using average room sizes and will likely utilize a range of average room sizes to better represent the dwelling areas of a region.

5 Discussion

This method's main limitation for future research is the time intensive process used to manually pinpoint and digitize the hundreds of individual rooms used to obtain the average room size. Ideally, this process would be able to expand to the entire country to create an Overall Residential population density number for all of Ghana, including the rural regions. Additionally, there is human error and bias associated with manually obtaining data points for digitization. We can address this by building onto one of our current PDT projects, which is a geospatial object detection tool, that will be trained to identify each instance of a waffle-home structure from the tiles of Digital Globe/Maxar imagery. This will create a data set of nearly every waffle home point in a country that can then be sampled in a similar manner as described in the methodology to find the average room size. This will ensure a wider spatial range of rooms identified to provide an average that better represents the wide spatial array of dwelling sizes. This image detection system can be easily implemented into existing imagery processing pipelines.

A secondary challenge involves the actual waffle home data points themselves. As seen from the literature, building priority has been placed on single-family and apartment style homes for a growing middle-class and less on compound style homes. As this study only samples buildings under construction from the past three years, there is a likelihood for the average room size to skew on the higher end. Samples were taken from more informal-type settlements and rural areas, but the majority were drawn from new urban construction. The use of the object detection system would increase the overall sample size of the waffle homes. Using this expanded sample, we could ensure a better balance in samples from rural and urban settlements towards a more representative sample of waffle homes. This, along with a future emphasis to breakout residential densities into socio-economic classes, as PDT already incorporates in its modeling, will be used to ensure lower-income and informal housing is better represented in these estimates.

This new methodology of utilizing the aerial imagery of rapidly expanding construction and GIS software to measure room areas has the potential to fill in data gaps for large patches of the world where normal data collection methods are unavailable. There is a need to model residential population density across the world, not only to be able to model as much of the world as possible, but also to better apply aid in humanitarian crises and environmental disasters. Unique data collection methods are necessary in areas where available data is lacking the necessary information. While this method has a fairly simple but unconventional approach, it is important to find new methods of data acquisition to circumnavigate challenges in data gaps.

References

1 Lewis Abedi Asante, Emmanuel Kofi Gavu, Jonathan Zinzi Ayitey, and Alexander Sasu. The changing face of compound houses in ghana and its effect on rental value: A case study of selected neighborhoods in kumasi, ghana. Technical report, African Real Estate Society (AfRES), 2015.

2 Samantha Duchscherer, Robert Stewart, and Marie Urban. revengc: An R package to Reverse Engineer Summarized Data. *The R Journal*, 10(2):114–123, 2018. doi:10.32614/RJ-2018-044.

3 Yinka Ibukun. How city life transformed ghana's compound houses. *Bloomberg: CityLab*, 2021. URL: https://www.bloomberg.com/news/features/2021-05-06/the-design-history-of-ghana-s-compound-houses.

4 Google Maps Imagery. Accra google maps imagery, 2023. Imagery @2023 Google, Imagery @2023 CNES / Airbus, Maxar Technologies, Map data @ 2023. URL: https://www.google.com/maps/place/Accra,+Ghana/@5.6250365,-0.1023598,198m/data=!3m1!1e3!4m6!3m5!1s0xfdf9084b2b7a773:0xbed14ed8650e2dd3!8m2!3d5.6037168!4d-0.1869644!16zL20vMGZueWM.

5 Iman Khajehzadeh and Brenda Vale. Estimating the floor area of a house knowing its number of rooms and how these are named. In *Back to the Future: The Next 50 Years, (51st International Conference of the Architectural Science Association (ANZAScA))*, December 2017.

6 Dahlia Nduom. Housing and culture in ghana: A model for research and evidence-based design. *ARCC Conference Repository*, September 2018. URL: https://www.arcc-repository.org/index.php/repository/article/view/533.

7 Ghana 2021 Population and Housing Census. Ghana 2021 population and housing census: Housing characteristics. Technical report, Ghana Statistical Service, 2021.

8 Ghana Statistical Service. Ghana living standards survey, report of the fifth round (glss 5). Technical report, Ghana Statistical Service, 2008.

9 Robert Stewart, Marie Urban, Samantha Duchscherer, Jason Kaufman, April Morton, Gautam Thakur, Jesse Piburn, and Jessica Moehl. A bayesian machine learning model for estimating building occupancy from open source data. *Natural Hazards*, 2016. doi:10.1007/s11069-016-2164-9.

10 Marie Urban, Robert Stewart, Scott Basford, Zachary Palmer, and Jason Kaufman. Estimating building occupancy: a machine learning system for day, night, and episodic events. *Natural Hazards*, 2023. URL: https://link.springer.com/article/10.1007/s11069-022-05772-3#citeas.