Targeting Water Subsidies Based on New Data Generation Technologies

4/24/2020

Objective:
To develop a tool, based on satellite imagery, google street view, and household survey data, which provides highly granular poverty maps of the urban landscape. An instrument to allow water utilities and regulators to better identify urban poverty, and permit for easier targeting of beneficiaries of water subsidies.

Lead Organization:

World Bank Water Global Practice (GP)

Lead Task Team Leader (TTL):

Camilo Lombana, Senior Water and Sanitation Specialist, Water GP, World Bank

Dr. Luis Alberto Andres – Lead Economist, Water GP, World Bank

Collaborators:

Dr. Gordon McCord - University of California San Diego School of Global Policy and Strategy

Dr. Marshall Burke – Stanford University Department of Earth System Science

Jonathan Grabinsky – Consultant, World Bank

Government of Angola Regulatory Institute for Water and Energy Services (IRSEA)

Instituto Nacional de Estatística de Angola (INE)

Global Water Challenge

Sustainable Development Goals (SDGs) covered:

SDG 1 – No Poverty

SDG 6 – Clean Water and Sanitation

SDG 10 – Reduced inequalities

SDG 16 – Peace, justice & strong institutions

Countries covered:

Angola

Data types and technologies:

Geospatial data

Official statistics

Household survey data

Satellite Imagery

Acknowledgements

This project, Targeting Water Subsidies Based on New Data Generation Technologies**,** submitted in response to the 2017 call for proposals by the World Bank's Development Data Group (DECDG) and the Global Partnership for Sustainable Development Data (GPSDD), is supported by the World Bank's Trust Fund for Statistical Capacity Building (TFSCB) with financing from the United Kingdom's Department for International Development (DFID), the Government of Korea, and the Department of Foreign Affairs and Trade of Ireland.

This work was co-led by Camilo Lombana, Senior Water and Sanitation Specialist of the Water GP at the World Bank, and Dr. Luis Alberto Andres, Lead Economist, Water GP, World Bank, in close collaboration with Jonathan Grabinsky, Consultant, World Bank, Dr. Gordon McCord, from the University of San Diego School of Global Policy and Strategy, Dr. Marshall Burke, from Stanford University Department of Earth System Science, and the Government of Angola's Regulatory Institute for Water and Energy Services (IRSEA). The Instituto Nacional de Estatística de Angola (INE) also provided assistance in the process. James Schull, Chenlin Meng, Cooper Ratenik, David Lobell, and Stefano Ermon from Stanford University provided support in analyzing the data, and piecing together the results. UDA Consulting (Turkey) and Cosep Consultora LDA (Angola) assisted with household-level data-collection.

Executive Summary

This study aimed to use remote sensing to generate highly granular poverty maps of Luanda. The poverty maps were meant to be used by Angola's Regulatory Institute for Water and Energy Services (IRSEA) to better allocate subsidies to those sectors of the population which need them the most. A better allocation of subsidies would help offset household's costs of connecting and paying for water services, thereby significantly expanding the percentage of the population with access to clean water.

The objective of this project therefore closely aligns with marking progress towards the achievement of Sustainable Development Goal (SDG) 6.1: "By 2030, achieve universal and equitable access to safe and affordable drinking water for all." Additional effects were expected in progress along other SDGs, particularly, in helping end poverty, in reducing inequalities (SDG 1 and 10, respectively), and in strengthening government institutions (SDG 16).[1]

The study combined sub-meter resolution DigitalGlobe (DG) imagery with data from 1,200 household surveys, and deployed a suite of classification algorithms to explore the correspondence between survey-based poverty measures and remotely sensed household information. The exercise sought to use high-resolution satellite imagery and street-view imagery to maximize the predictive power of satellite-based measures using two continuous measures of poverty: 1) the household's aggregate monthly income and 2) a multidimensional poverty index (MPI).

Using state-of-the-art machine learning algorithms, findings suggest that the model had relatively low levels of predictive accuracy, insufficient to provide IRSEA with any dependable mapping of urban poverty in Luanda. A city-wide prediction using nighttime and daytime imagery on an additional measure of household-wealth – a combined Wealth Index (WI), built from a principal-component analysis (PCA) of household infrastructure and assets – is provided as proof-of-concept of how these models could be used for targeting poor households for public support programs. Nevertheless, the goodness-of-fit estimates of this proof-of-concept exercise remain low. Overall, the project was unable to produce a reliable, high-definition, poverty prediction map of Luanda.

A series of lessons learned which could help improve the precision of this remote sensing model moving forward include: a) further refine household survey data-collection in order to improve the accuracy of georeferenced data, b) improve the street imagery captured by the surveyors, c) test the machine-learning algorithms model using alternative indicators of poverty, d) apply the model to different cities, in the hope that these are more responsive to high-quality satellite and street-view imagery, and can thus be more easily calibrated in the remote sensing exercise.

[1] United Nations Sustainable Development Goals. Goal 6: Ensure access to water and sanitation for all https://www.un.org/sustainabledevelopment/water-and-sanitation/. Date Accessed: April 16, 2020

Project Background

A significant portion of the population in low- and medium-income countries are unable to afford the high costs of paying for water services. To help offset the costs, governments often offer subsidies. But absent granular, accurate, data on the poverty distribution of families within a country, subsidies are frequently allocated in a sub-optimal way; monetary assistance often goes to wealthy and poor families alike, even if well-off families are not in need of financial support. And poor households, which are more urgently in need of access to subsidies to cover their service bills, are, even with government support, often left with insufficient funds to pay for their ongoing water expenses.

With the aim of improving the allocation of water subsidies, the World Bank Water Global Practice (GP), in collaboration with academics from the University of California San Diego School of Global Policy and Strategy and Stanford University Department of Earth System Science, sought to develop a model to create predictive urban poverty maps. The initiative, titled Targeting Water Subsidies Based on New Data Generation Technologies was piloted in Luanda, Angola. If successful, the aim was for the tool to be scaled-up across Angola and other countries, to help optimize the allocation of water subsidies.

The central objective of this project was to create a poverty-prediction tool which could be used by IRSEA to better target its water subsidies. Subsidies which could help offset household's fixed and variable costs of connecting to water services, thereby significantly expanding the population with access to clean water. In this way, the initiative is strongly aligned with SDG Target 6.1: " by 2030, achieve universal and equitable access to safe and affordable drinking water for all." In helping improve the government's allocation of resources, additional effects were also expected across other domains: in alleviating the economic conditions of households (SDG 1 and 10), and in bettering IRSEA's institutional capacity (SDG 16).

A state-of-the-art object detection algorithm, called You-Only-Look-Once (YOLO), was used to extract counts of objects and features from 60 predefined objects, including buildings, small cars, houses, and trucks. Uni-modal and deep multi-modal methods were then used to extract features from image data and predict poverty measures from these features. The predictions were applied using two different metrics of poverty: aggregate monthly household income and a multidimensional poverty indicator. Findings suggest that the model had relatively low levels of predictive accuracy, insufficient to provide IRSEA with any dependable citywide map of urban poverty in Luanda.

A city-wide prediction on a WI – built via a PCA of household infrastructure and assets – using medium-resolution nighttime and daytime imagery is provided as proof-of-concept of how these models could be used to target poor households for public support programs. However, overall, the goodness-of-fit of these models remains low; insufficient to produce poverty maps which are useful for policy purposes.

Lessons learned and next steps to help improve this exercise include: a) refine household survey data-collection in order to improve the accuracy of georeferenced data, b) improve the street imagery captured by the surveyors, c) test the machine-learning, remote-sensing model using

alternative indicators of poverty, d) and apply the model to different cities, in the hope that these respond better to higher-quality satellite and street-view imagery, and can thus be more easily calibrated in the machine-learning models.

This report follows the subsequent structure: the next section covers the data & methodology used for the exercise, followed by the results of the modeling exercise, the limitations of the data, the risks faced, pinpointed responses to the key indicator output questions, the lessons learned from the process, and the conclusion.

Data & Methodology

The model was piloted in Luanda, and used Remote Sensing Imagery consisting of daytime, nightlights, and radar imagery. It was calibrated using household surveys and images from 1,200 households from Luanda, and added geospatial datasets including highways, roads, schools, and public health facilities, as covariates. The YOLO model used here extracts counts of objects in DG imagery across 60 predefined classes, including buildings, small cars, buses, and trucks.

Uni-modal and deep multi-model methods were then used to extract features from the image data and predict poverty measure using two different outcomes: aggregate household monthly income and a MPI. The results suggest that the model had relatively low levels of predictive accuracy, insufficient to provide IRSEA with a reliable citywide map of urban poverty in Luanda.

Alternatively, a city-wide prediction on an additional measure of household-wealth – using a combined WI built from a PCA of household infrastructure and assets – with medium-resolution nighttime and daytime imagery is provided as proof-of-concept of how these models could be used for targeting poor households for public support programs.

This section offers an overview of the different datasets used, the household poverty indicators built, as well as a rundown of the methodology behind the machine-learning exercise.

Data

A non-representative household survey data was collected during February 2019. A total of 1,200 households were selected across the city in the following way: the city of Luanda was divided into 13,764 grid cells of 250 by 250 meters each. The sample frame was limited to the 5,848 grid cells that were not classified as having "very low poverty" and limited those that were within one-half a standard deviation from the population mean for the city.

Each grid cell was assigned a poverty classification using a visual mapping from existing poverty maps for the city[2], and population for each grid was calculated using European Commission 2019 GHSL - Global Human Settlement Layer. [3] From the 5,848 grids, 240 grids were randomly

[2] Development Workshop Angola (2011). "Poverty and Environmental Vulnerability in Angola's Growing Slums: Comparative Analysis of Luanda, Huambo and Cachiungo." Working paper.
[3] European Commission (2019). GHSL - Global Human Settlement Layer. Available at: https://ghsl.jrc.ec.europa.eu/

selected with numbers selected from each remaining poverty stratum to match the actual population across the poverty strata[4]. Five households were randomly sampled from each selected grid, for a total of 1,200 households across the city[5]. Figure 1 shows the sampling frame grids and sample households in Luanda.

Two household indicators were constructed for use in the machine-learning classification exercise. The first is monthly-level household income, aggregated across all members of the family. A second household indicator for analysis was a weighted, household-level deprivations index, following the MPI approach recommended by the Oxford Poverty & Human Development Initiative Global Multidimensional Poverty Index[6] . The different components that were fed into the MPI, and their respective weights, are included in Table 1.

Given the low predictive capacity of the machine-learning classification exercise, a city-wide prediction mapping exercise, using nighttime and daytime imagery, was administered as proof-of-concept. The prediction mapping exercise was conducted on an aggregate WI measure: the WI is a composite measure of a household's wealth, built by running a PCA on a household's ownership of selected household infrastructure and assets: such as source of drinking water, type of toilet, material of principal floor and household assets. Table 2 shows the different indicators that were fed into the PCA to calculate the WI. [7]

---

[4] After eliminating cells with "Very Low" poverty, the remaining poverty strata in Development Workshop Angola (2011) are "Very High," "High," "Moderate," and "Low". From these, 186 cells were chosen from the "Very High" poverty cells, and 35, 10 and 9 from the others, respectively. This stratification weighted cell selection according to the actual population distribution across these poverty strata. Note that, once in the field, 34 of the originally selected grids were found to be not residential, or to have fewer than 15 buildings. These 34 were replaced with random selection from the list of remaining grids in the same poverty stratum as the grid lost.

[5] IRB approval for the project was secured from the University of California, San Diego.

[6] Alkire, Sabina and Selim Jahan. "The New Global MPI 2018: Aligning with the Sustainable Development Goals," HDRO Occasional Paper, September 2018.

[7] Most of these factors were also used to build the wealth index for the Angola Demographic Household Survey (DHS) 2015. See: The DHS Program - Research Topics - Wealth Index. https://dhsprogram.com/topics/wealth-index/. April 23, 2020

Figure 1: Sampling Frame Grids and Sampled Households in Luanda, Angola

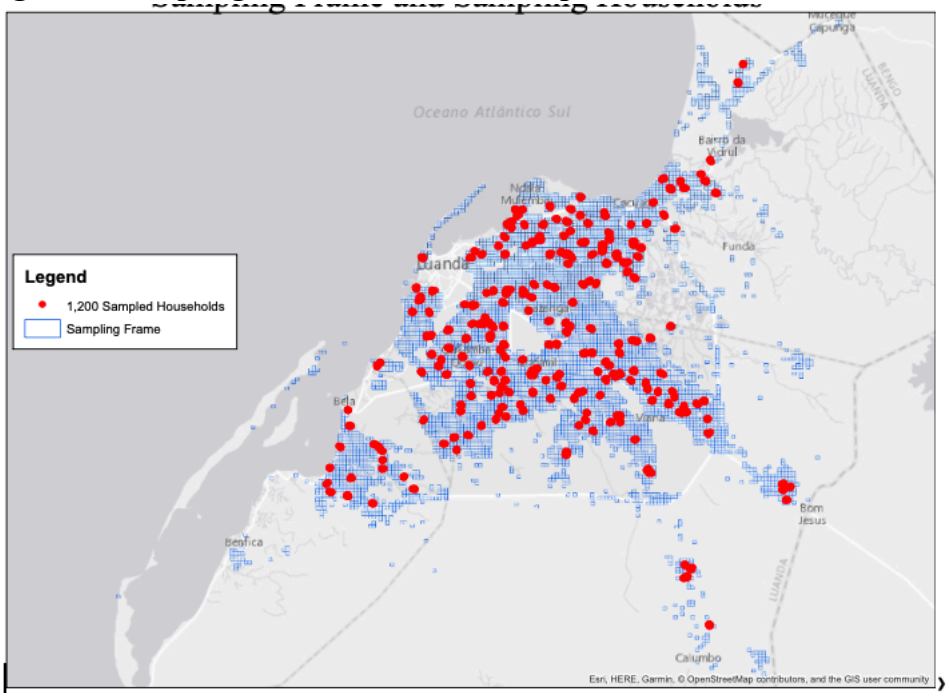


Table 1) The MPI indicators and respective weights

Dimensions of Poverty	Indicator	Deprived if Living in the Household Where	Weight
Health	Nutrition	Any adult under 70 years of age or a child is undernourished	1/6
	Child Mortality	Any child under the age of 18 years has died in the five years preceding the survey	1/6
Education	Years of Schooling	No household member aged 10 years or older has completed six years of schooling	1/6
	School Attendance	Any school-aged child is not attending school up to the age at which he/she would complete class 8	1/6
Standard of Living	Cooking fuel	The households cooks with dung, wood, charcoal or coal	1/18
	Sanitation	The household's sanitation facility is not improved (according to SDG guidelines) or it is improved but shared with other households	1/18
	Drinking water	The household does not have access to improved drinking water	1/18
	Electricity	The household has no electricity	1/18
	Housing	Housing materials for at least one of roof, walls and floor are inadequate: the floor is made of natural materials and/or the roof and/or walls are of natural or rudimentary materials	1/18
	Assets	The household does not own more than one of these assets: radio, TV, telephone computer, animal cart, bicycle, motorbike or refrigerator, and does not own a car or truck.	1/18

Table 2) Indicators Included loaded onto the PCA to develop the WI

Indicators used to build the Wealth Index		
People Per Sleeping Room	Main Source of Light Used	Cell Phone
Main Source of Drinking Water	Radio	Bicycle
Location of Water Source	T.V.	Car
Type of Sanitation Facility	Fixed Telephone	Motorboat
Main Power or Fuel Source Used by The Household	Computer	Animal Traction Cart
Main Source of Illumination Used	Internet	Floor Material
Wall Material	Roof Materials	

Note: all variables without variance are excluded. PCA generates factor loadings as weights, which allow us to create an index variable at the household-level. Most of these factors were also used to build the wealth index for the Angola Demographic Household Survey (DHS) 2015

For each household in the survey dataset, enumerators photographed the front of the household as well as the surrounding street environment. Each household had high-resolution street view images at the precise survey locations, for a total of 7,200 street view images in total. Figure 2 displays a series of examples of the images taken by the enumerators on the ground. These street- view images were used to help calibrate the model.

Figure 2: Randomly selected street view imagery

To enrich the predictive capacity of the model, publicly provided infrastructure and services including highways and roads, schools, and public health facilities were added as covariates to the model. Images of the covariates used are included in the appendix. Distances were calculated from each household in the sample to the nearest of each type of facility and road type, and included as

potential predictors. This was done under the hypothesis that a statistically significant relationship exists between the proximity and quality of public goods and a household's income and MPI.[8]

For each household, a 0.3m resolution DG imagery was used (yielding 1000x1000 pixel images), 500m resolution VIIRS Nightlights imagery using the median of 2018 VIIRS nightlights data (yielding 225x225 pixel images), and 10m resolution Sentinel-1 Radar imagery using the median of 2018 Sentinel-1 imagery (yielding 225x225 pixel images). The daytime, nightlights, and radar imagery was preprocessed and transformed to extract different features according to various models.

Methodology

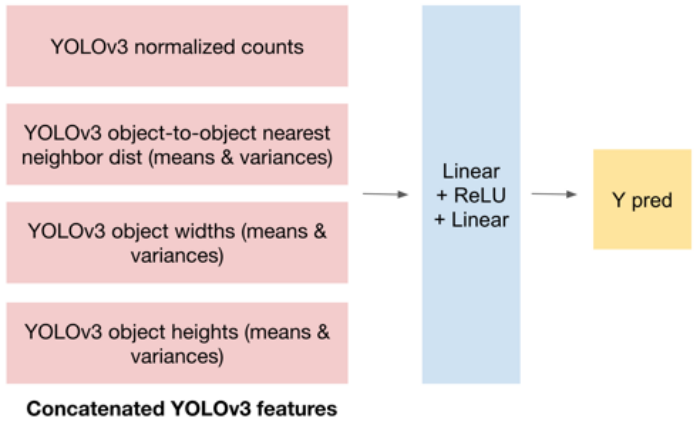
The specific model used was YOLO version3, a state-of-the-art object detection algorithm. YOLO extract counts of objects in DG imagery across 60 predefined classes, including buildings, small cars, buses, and trucks. Each tile's outputs are the bounding boxes of detected objects and their corresponding confidences. Each of the satellite imagery sources is centered at the exact latitude and longitude of each household. The object counts in each image tile are then used a predictor of the income or deprivations outcome variable.

The model uses the following three features as predictors of poverty: normalized image counts, nearest neighbor distances and means, and the widths and weights of the objects (Figure 3).

- Normalized image counts take the raw object counts in a sample and normalize it within the sample. For example, an image with 3 cars and 2 buildings will be normalized to $3/5$ cars and $2/5$ buildings.
- Nearest neighbor distances and means calculate the distance to the closest neighbor in each other object category. This, in turn, is an estimate of the aggregate density of objects in a region — the presumption is that the density of buildings or cars can serve as a proxy indicator for socioeconomic conditions.
- Lastly, the width and weights of objects are measured, and they are often assumed to serve as indicators for poorer or wealthier area: larger buildings can indicate spaciousness while smaller buildings can indicate overcrowded living conditions.

[8] Highway and school data were taken from OpenStreetMap (OMS), a web-based, collaborative mapping projects that allows users to freely edit maps with geographic information inputs. Public-health facilities data come from two sources: publicly available geocoded inventory of public hospitals from 48 sub-Saharan countries collected from Ouma, P.O. et al. 2018, and data obtained from the World Bank from GEPE/MINSA (Gabinete do Estudo e Planeamento e Estatística/Ministério da Saúde).

Figure 3: YOLOv3 extracted features through a 2-layer Neural Network



Uni-model and deep multi-model methods were used to extract features from the image data and predict poverty measures from these features. Each of the methods was applied to the poverty index using both the income regression and the MPI. All models developed in this paper used Mean Squared Error (MSE) as their loss function. Moreover, to serve as an additional cross-check ,two people were trained to look at 100 random samples of the data to test the probability of them accurately detecting the level of the deprivation of the household. (Table 4).

Given the low-predictive capacity of the models mentioned under the "Results" section below, a prediction mapping exercise, using medium-resolution nighttime and daytime imagery, was administered as proof-of-concept of how these models could be used to target poor households for public support programs. However, it should be stressed that these exercises are only to be used as a proof-of-concept, and that the goodness-of-fit is not high enough for the poverty maps to be used to inform policy.

Results

The 1,200 household data points were randomly shuffled, 75% of the data was used as the training set and the remaining 25% as the test set. This resulted in 900 training samples and 300 test samples.

Table 3 present the results of the machine-learning exercise on the income outcome variable, and on Table 4 presents it on the MPI. Table 3 depicts the performance of the models, measured by the coefficient of determination r^2 between the test set predictions and ground truth, in regression on household income. None of these models are working particularly well, as the R-squared never rises above 0.12.

Table 4 depicts the performance of the models, measured by the accuracy of predictions on the test set, in classification on the MPI. The models are indeed able to learn some small amount of predictive signal from the training data, achieving a higher accuracy than the human baseline.

The low r^2 depicted in Table 3 and the low levels of accuracy depicted in Table 4 suggests that there is a large degree of noise in the models, as they are unable to generalize from the training set to the test set, and that the goodness-of-fit of these models is not high enough to be able to generate citywide poverty maps.

Table 3) Model Results on Income

Model results (income)		
Model	Features	**r^2 (test)**
Linear regression	Street-view with color histogram	**0.043**
Linear regression	YOLO	**0.032**
MLP	Flattened Digital Globe images	**0.005**
MLP	DenseNet embeddings of Digital Globe images	**0.012**
MLP	DenseNet embeddings of Sentinel 1 images	**0.02**
Multimodal CNN	**All inputs**	**0.12**

Table 4) Model Results on the MPI.

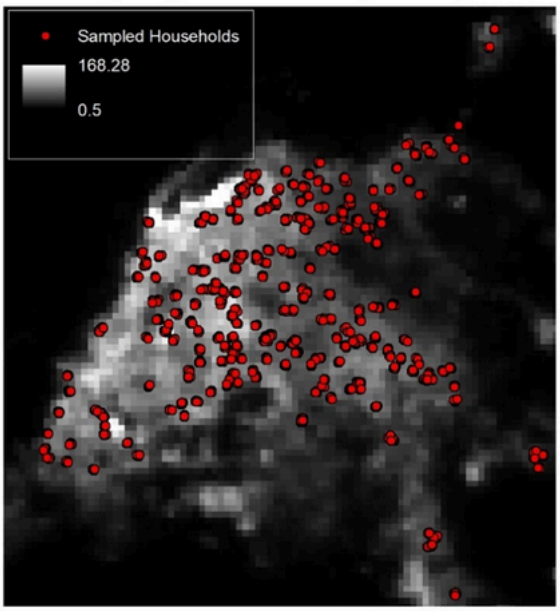
Model results (MPI)		
Model	Features	**Accuracy (test)**
Human	All	**49.5%**
Predict-1	None	**54.1%**
Multi-modal CNN	Street-view + YOLO	**56%**
Multi-modal CNN	Street-view + YOLO + DG + S1/VIIRS	**56%**
Multi-modal CNN (2-layer MLP)	Street-view + YOLO + DG + S1/VIIRS	**57%**
Multi-modal CNN (1-layer MLP)	Street-view + YOLO + DG + S1/VIIRS	**52%**
Multi-modal CNN	**Street-view**	**58%**

Despite the relatively low accuracy of prediction using high-resolution imagery and the methods described above, a city-wide prediction mapping exercise was conducted as proof-of-concept. Since neither DG imagery nor street view photographs are available for all households in the

sampling frame, this exercise was conducted using publicly available medium-resolution city-wide imagery.

Landsat 8 imagery was used to provide a cloud-free image of the city matching the survey date in February 2019, and to construct the following urban landcover indices: The Normalized Difference Built-Up Index (NDBI), The Urban Index (UI), and The Enhanced Built-Up and Bareness Index (EBBI). Also used was nighttime imagery from the VIIRS satellite, which provides nighttime radiance at a spatial resolution of 500 meters. The VIIRS satellite model uses the median value for each pixel between January 1 and March 31, 2019. Figure 5 shows the variation in nightlights in Luanda during the survey period, together with the location of the sampled households. The spatial geographic covariates mentioned under the "Data & Methodology" section, and referenced in the appendix, were also included in the model.

Figure 5: Nightlights Radiance and Sampled Households Location



The city-wide prediction mapping exercise was applied on the WI variable. This is because, when running a model selection exercise across different outcome indicators, WI holds the most predictive power (see Table 6 in the Appendix). A stepwise regression was first conducted on the WI, using all available predictors, keeping only variables with p-values above 0.05 (Table 5).

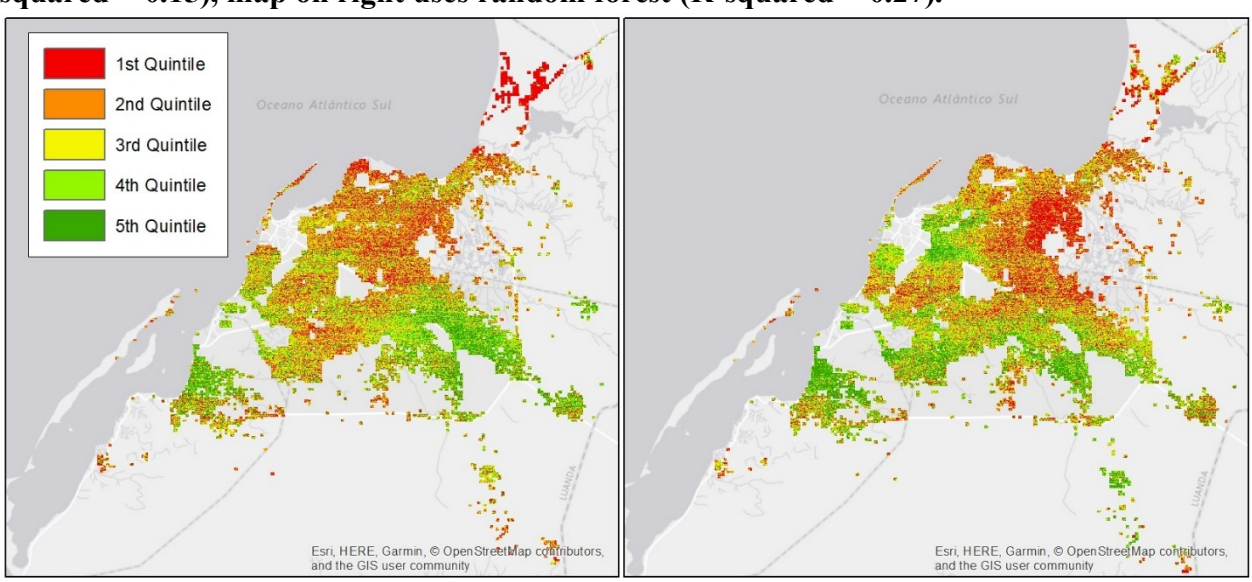
Table 5) Stepwise Regression of WI on all available predictors

Independent Variables	Coefficient	t-statistic
Nightlight radiance	0.0322	7.79
Distance to hospital	0.000078	4.16
Distance to school	0.000072	2.24
Distance to secondary road	-0.00012	-2.73
Distance to tertiary road	-0.00029	-2.74
NDBI	-3.93	-2.81
UI	3.75	2.43
Landsat 8 Band 1	130.4	4.09
Landsat 8 Band 2	-101.3	-2.59
Landsat 8 Band 3	-34.87	-2.43
Landsat 8 Band 6	16.24	4.80
Landsat 8 Band 10	-0.318	-2.82
Landsat 8 Band 11	0.255	2.36
Number of observations	1,200	
R-squared	0.15	

Note: keeping only variables with p-values above 0.05.

The values for each variable in the model above were applied to the entire sampling frame of the study at a spatial resolution of 10-meters (3.9 million points in total). The value of the WI was calculated at each location using both the linear regression model and the random forest model, and the predicted values of WI were then graphed. See Figure 4 below.

Figure 4) WI Prediction. Map on left uses predicts using linear regression model (R-squared = 0.15), map on right uses random forest (R-squared = 0.27).

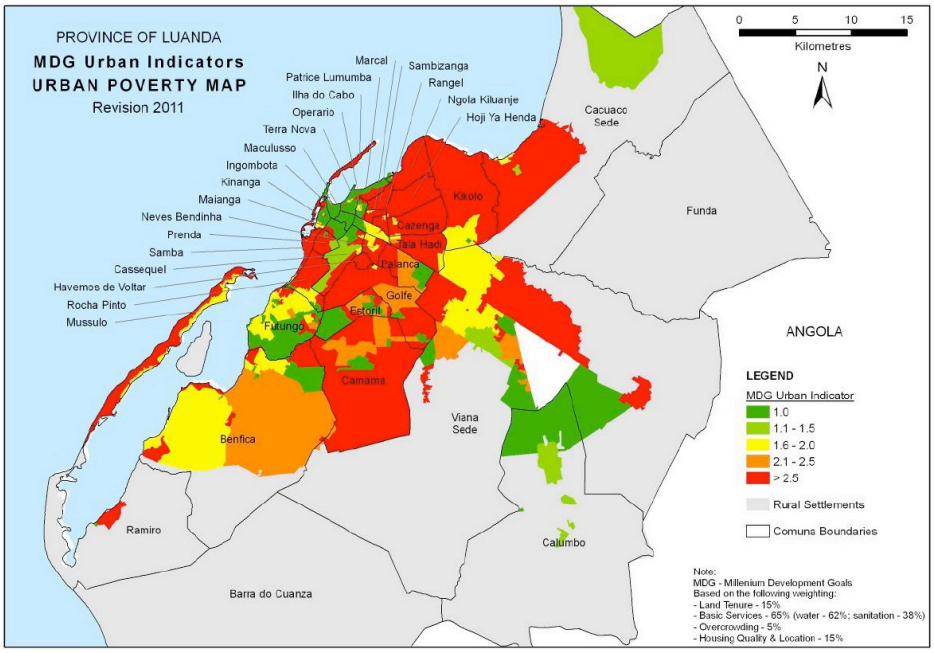


Given the low goodness-of-fit attained in both the regression and random forest models, the maps above should be interpreted with caution: these exercises should be understood as proof-of-

concept, and not a final product for IRSEA, or other policymakers, to use to infer citywide household wealth.

Nevertheless, on a positive note, the poverty distribution displayed in the maps more or less coincide with a previously published poverty map of Luanda generated by Development Workshop Angola (2011)[9]. As mentioned under the "Data" section of this report, the poverty distribution from this map, displayed in Figure 5, was used as a reference to partition Luanda into the 13,764 grids that were used for sampling of the household survey.

Figure 5) Urban Poverty Map for Luanda, from Development Workshop Angola (2011)



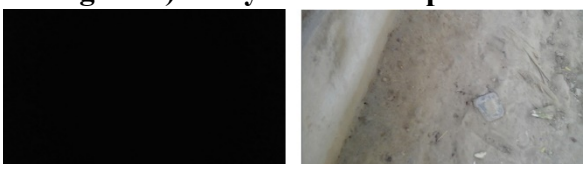
Data Limitations:

As mentioned previously, the model proved unable to generate high-definition poverty estimates of Luanda. There were a considerable number of limitations in the household dataset, and the DG/nightlight imagery**,** that hindered the model's predictive capacity. These limitations include: noise in the street-view of the household dataset, seeming uncorrelated inputs and labels in the household dataset, and satellite imagery overlap.

a) **Noisy street-view in household dataset**: A large proportion of the street-view images were noisy, either containing no information at all, or containing unusable information (e.g. all six photographs are simply of the floor). During a pass of 50 random training examples, street-view samples for 3 households contained no information, implying around 6% of the whole dataset has noisy street- view imagery.

[9] Development Workshop Angola (2011). "Poverty and Environmental Vulnerability in Angola's Growing Slums: Comparative Analysis of Luanda, Huambo and Cachiungo." Working paper.

Figure 6) Noisy street-view photos



b) **Seemingly uncorrelated inputs and labels in household dataset**: Some of the ground truth input-label pairs were very surprising to a human observer. Some of the most visibly wealthy homes in the dataset (in terms of size, neighborhood, cars), for example, had relatively high deprivations scores.

Figure 7) Comparison between households with minimum and maximum deprivation scores (amongst examples in the datasets

c) **DG /nightlight imagery overlap**: Satellite image overlap, resulting from the resolution of the imagery and the proximity of households, caused households with vastly different income and deprivations indices to be associated to near identical DG and Sentinel 1/VIIRS images. This suggests that the sampling strategy and/or the urban landscape of Luanda may have proved too difficult a classification for remote sensing.

Risk Mitigation

The household survey fieldwork and the building of the machine-learning model unfolded without any noteworthy risks. IRSEA, and, to a lesser extent, INE, were involved throughout the process in seeing the project succeed.

Output Indicator Questions

In order to guarantee compliance with the funding requisites from the TFSCB, this section offers detailed responses to the following output-indicator questions:

- Have there been any final results or outcomes in which data or methods have allowed data to be produced: faster; more cheaply; at a higher resolution or granularity, or where there was no data before?
- Has the project contributed to the production and/or use of data disaggregated by a) sex b) disability c) age, d) geography (or other)?
- Has the project contributed to the use and/or production of gender statistics? If yes, please describe.

Have there been any final results or outcomes in which data or methods have allowed data to be produced: faster; more cheaply; at a higher resolution or granularity, or where there was no data before?

As indicated previously, the findings suggest that the model had relatively low levels of predictive accuracy, insufficient to provide IRSEA with any dependable citywide map of urban poverty in Luanda.

The city-wide nighttime imagery map produced in Figure 4, used as a proof-of-concept, was successful in producing poverty data at a high-level of geographic detail, where no data existed. However, given the overall low goodness-of-fit across the models, the estimates can't be generalized or used to develop reliable citywide poverty-estimates.

Additional work is needed to provide maps, and a poverty-tool, which can be used by the government to better target water subsidies.

Has the project contributed to the production and/or use of data disaggregated by a) sex b) disability c) age, d) geography (or other)?

This remote-sensing exercise was aimed at producing citywide predicted poverty maps. In this way, the proof-of-concept exercise outlined above produced estimates with very high levels of geographic detail. However, the goodness-of-fit across the models is so low that the estimates can't be used to develop reliable citywide poverty estimates.

The aim of the project was not to provide data disaggregated by sex, disability, age or other socioeconomic variables. However, the information from the non-representative household survey can be used to obtain a (non-representative) snapshot of some of these patterns in Luanda.

Has the project contributed to the use and/or production of gender statistics? If yes, please describe.

The aim of the project was not to provide data disaggregated by gender. The focus was primarily on providing poverty estimates. However, the information from the non-representative household survey can be used to obtain a (non-representative) snapshot of some of these patterns, by gender.

Lessons Learned

This initiative has shed light on some of the limitations behind remote-sensing exercises. A series of steps to help improve these exercises in the future include: a) refine approaches to improving household-level data-collection, with a special focus on improving the georeferencing capacity of the devices used b) improve the street imagery captured by the surveys, c) test the machine-learning algorithms model using alternative indicators of poverty, and d) apply similar remote-sensing models across different cities, in the hope that alternative urban landscapes may respond better to satellite and street-view imagery.

Significant steps were taken to ensure that the survey data collected was of high-quality. However, additional work is needed to further improve household-level data collection. This is especially true when it comes to improving the georefencing capacity of the devices used, as the remote sensing exercise is deeply sensitive to geographic precision. Any future work should thus aim to ensure that the georeferencing devices being used by the surveyors are of the highest geographic precision possible.

Moreover, additional work is needed to improve the street imagery captured by the surveys. While promising, the street view imagery collected by enumerators were extremely variable in quality, orientation of picture, lighting, and other characteristics (see Figure 7). This lack of uniformity compromised the ability of algorithms to detect patterns usable for prediction.

The remote-sensing model should also be tested using alternative indicators of poverty. The two indicators used as part of the remote sensing exercise where: the household's aggregate monthly income and the MPI. However, additional indicators of poverty, using aggregate monthly household consumption instead of income, and applying the WI more broadly – beyond the proof-of-concept nighttime and daytime imagery city-wide prediction to the remote-sensing exercise itself – could potentially produce more reliable estimates.

Lastly, it's important to test these remote-sensing models across different urban landscapes. It may be that the selected sample proved too hard a classification exercise with remote sensing, and that the sampling frame in this project focused on too heterogeneous neighborhoods. This raises the need for further research and testing need to be conducted in other settings, and across other samples.

Conclusion

The study combined sub-meter resolution DG imagery with data from 1,200 household surveys, and deployed a suite of classification algorithms to explore the correspondence between survey-based poverty measures and remotely sensed household information. The exercise sought to use high-resolution satellite imagery and street-view imagery to maximize the predictive power of satellite-based measures using two continuous measures of poverty: 1) the household's aggregate monthly income and 2) a MPI.

Findings suggest that the model had relatively low levels of predictive accuracy. A city-wide prediction using nighttime and daytime imagery on an additional measure of household-wealth – a combined WI, built from a PCA of household infrastructure and assets –is provided as proof-of-concept of how these models could be used for targeting poor households for public support programs. Nevertheless, the goodness-of-fit estimates of this proof-of-concept exercise remain low. Overall, the project was unable to produce a reliable, high-definition, poverty prediction map of Luanda.

A series of next steps to be taken to improve these exercises in the future include: a) refine approaches to household-level data-collection, with a special focus on improving the georeferencing precision of the devices used b) improve the street imagery captured by the surveys, c) test the machine-learning algorithms model using alternative indicators of poverty, and d) apply similar machine-learning models across different cities, in the hope that other urban landscapes may respond better to satellite and street-view imagery.

This exercise attempted to provide the government with highly granular poverty information which could be used by IRSEA to better allocate its resources. Although the models tested had very low-levels of predictive accuracy, this line of remote-sensing work holds promise in helping governments better pinpoint poverty. The hope is that similar exercises can be further replicated within the WASH sector, to allow government to identify those households that need subsidies the most, and, in line with SDG 6.1 and 6.2, help thread progress towards achieving universal access to WASH by 2030.

Appendix

Table 6) Model selection exercise across different outcome indicators for the goodness-of-fit exercise using Random Forest and Lasso.

Outcome	Model	Goodness-of-fit
Poor (deprivations index > 0.33)	Random forest	Balanced accuracy = 55%
Severely Poor (deprivations index > 0.5)	Random forest	Balanced accuracy = 50%
Access to clean water (1/0)	Random forest	Balanced accuracy = 63%
Deprivations Index	Random forest	R-squared = 0.05
	Linear LASSO	R-squared = 0.03
Wealth Index	Random forest	R-squared = 0.27
	Linear LASSO	R-squared = 0.12
Income	Random forest	R-squared = 0.04
	Linear LASSO	R-squared = 0.004

Maps of Covariates

Figure 7) Primary Road Network in Luanda

Major Highway Systems, by Road Type

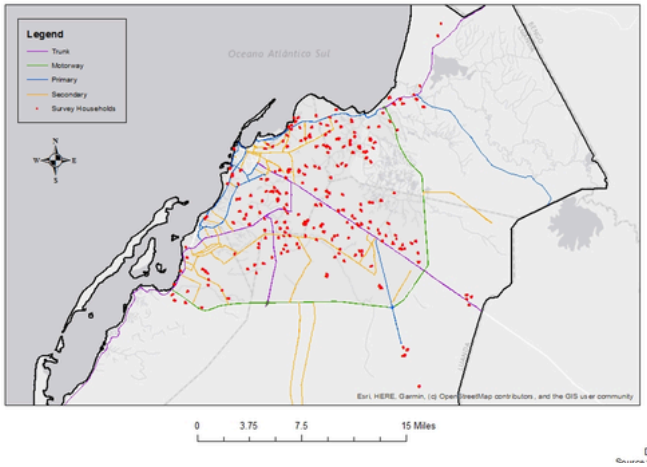


Figure 8) Paved Road Network in Luanda

Paved Roads and Survey Households

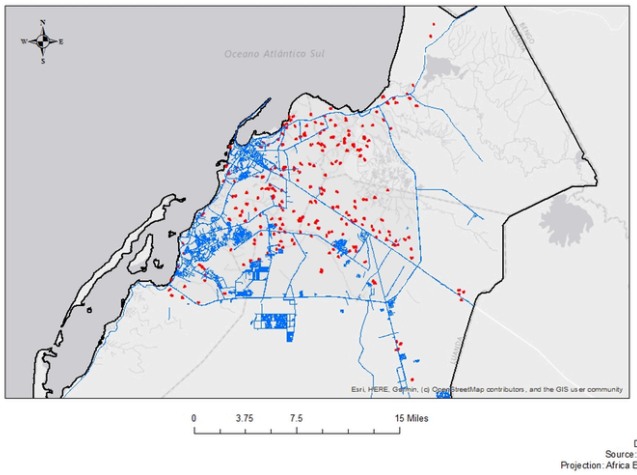


Figure 9) Spatial Distribution of Public Services Luanda

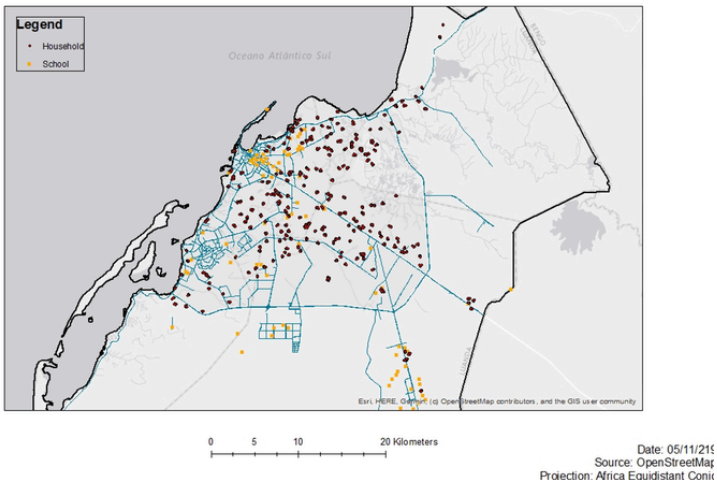


Figure 10) Government Hospitals in Luanda

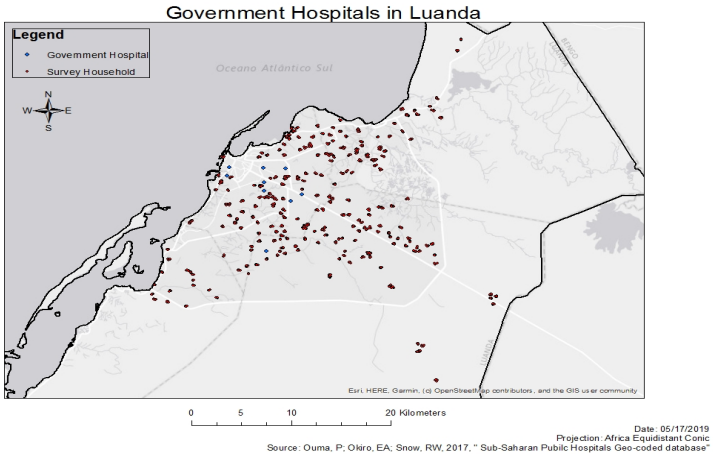
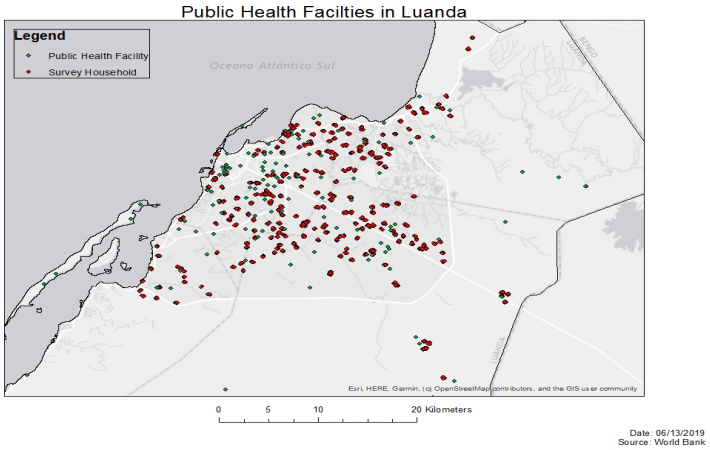


Figure 11) Public Health Facilities in Luanda



Work Cited

Alkire, Sabina and Selim Jahan. "The New Global MPI 2018: Aligning with the Sustainable Development Goals," HDRO Occasional Paper, September 2018.

As-syakur, Abd. Rahman et al. (2012) "Enhanced Built-Up and Bareness Index (EBBI) for Mapping Built-Up and Bare Land in an Urban Area." *Remote Sensing* 4:2957-2970.

Caruana, Rich and Steve Lawrence, and C Lee Giles (2001). "Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping," in *Advances in neural information processing systems*, pages 402–408.

Development Workshop Angola (2011). "Poverty and Environmental Vulnerability in Angola's Growing Slums: Comparative Analysis of Luanda, Huambo and Cachiungo." Working paper.

European Commission (2019). GHSL - Global Human Settlement Layer. Available at: https://ghsl.jrc.ec.europa.eu/

Gebru, Timnit, Jonathan Krause, Yilun Wang, Duyun Chen, Jia Deng, Erez Lieberman Aiden, and Li Fei-Fei (2017). "Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States." *Proceedings of the National Academy of Sciences*, 114(50):13108–13113.

He, Kaiming and Ross Girshick, and Piotr Dollar (2019). "Rethinking imagenet pre-training," in *Proceedings of the IEEE International Conference on Computer Vision*, pages 4918–4927.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil. "Measuring economic growth from outer space." American economic review 102.2 (2012): 994-1028.

Jean, N. et al. "Combining satellite imagery and machine learning to predict poverty." *Science* vol. 353, 6301: 790-794. August 19, 2016. doi: 10.1126/science.aaf7894.

Kingma, Diederik P. and Jimmy Ba (2014). "Adam: A method for stochastic optimization."

M. Xie, N. Jean, M. Burke, D. Lobell, S. Ermon, "Transfer learning from deep features for remote sensing and poverty mapping," AAAI Conference on Artificial Intelligence (2016).

Ouma, P.O. et al. "Access to emergency hospital care provided by the public sector in sub-Saharan Africa in 2015: a geocoded inventory and spatial analysis." *The Lancet. Global health* vol. 6,3 (): e342-e350. doi:10.1016/S2214-109X(17)30488-6.

Reddi, Sashank J and Satyen Kale, and Sanjiv Kumar (2019). On the convergence of adam and beyond. arXiv preprint arXiv:1904.09237.

Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi (2016). "You only look once: Unified, real-time object detection." in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788.

The DHS Program - Research Topics - Wealth Index. https://dhsprogram.com/topics/wealth-index/. Date Accessed: April 16, 2020.

United Nations Sustainable Development Goals. Goal 6: Ensure access to water and sanitation for all https://www.un.org/sustainabledevelopment/water-and-sanitation/. Date Accessed: April 16, 2020.

Zhihuan, Wu, Chen Xiangning, Gao Yongming and Li Yuntao. "Rapid Target Detection in High Resolution Remote Sensing Images Using YOLO Model," The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-3, 2018.