

# Global poverty estimation using private and public sector big data sources

Robert Marty<sup>1</sup> and Alice Duhaut<sup>\*1</sup>

<sup>1</sup>World Bank

August 18, 2023

\*Corresponding author: [adhaut@worldbank.org](mailto:adhaut@worldbank.org)

## Abstract

Household surveys give a precise estimate of poverty; however, surveys are costly and are fielded infrequently. We demonstrate the importance of jointly using multiple public and private sector data sources to improve predictions of levels and but also changes in wealth for a large set of countries. We train models using 63,854 survey cluster locations across 59 countries, relying on data from satellites, Facebook Marketing information, and OpenStreetMaps. The model generalize previous studies to a wide set of countries and explains 61% of the variation in levels of wealth at the survey cluster level and 68% of the variation at the district level, and the model explains 5% and 10% of the variation of changes in wealth at the cluster and district levels. Models perform best in lower income countries and in countries with higher variance in levels and changes in wealth. Features from OpenStreetMaps, nighttime lights, and land cover data are most important in explaining levels of wealth, and features from nighttime lights are most important in explaining changes in wealth.

# 1 Introduction

Accurate estimates of poverty are important for the design, delivery and evaluation of social programs. However, obtaining accurate estimates of poverty through household surveys is expensive and time consuming. Many countries conduct just a couple of nationally representative survey with poverty estimates within the span of a decade, resulting in countries relying on outdated estimates [30]. For example, among low and lower-income countries in the World Bank poverty calculator database, the latest census is 5.5 years old and the latest nationally representative household survey is 4.5 years old [3]. Moreover, few survey clusters are surveyed multiple times, making it difficult to track changes in wealth or poverty over time at a granular level [48].

To generate cheap, timely, and accurate poverty estimates, a growing literature has leveraged the emerging availability of global, spatially-referenced datasets to estimate poverty [6]. Socioeconomic surveys are used to train machine learning models that rely on features derived from these data sources so that poverty estimates can be extrapolated across time and space. Much of the work has focused on the use of nighttime and daytime satellite imagery [30, 48]; nighttime lights are a widely used proxy for local economic activity [11], and daytime imagery can capture features such as vegetation, built-up areas, roads—and the spatial configuration of these features. Additional research has shown the value of other data sources for poverty estimation, including anonymized cell phone data records (CDR); global human settlement data; Facebook marketing data (for example, capturing the percent of Facebook users with an expensive phone); OpenStreetMap data to understand accessibility to services like schools and health centers; and satellite-derived features beyond nighttime and daytime imagery including climate and environmental features, land use, and physical attributes such as elevation [38, 18, 37, 40, 28, 50, 36].

With the increasing availability of global, spatially-referenced data sources, it is useful to test the value added of bringing multiple datasets together and the contexts under which poverty estimation works best. We answer the following three related research questions. First, to what extent can multiple data sources be used to estimate not only levels but also changes in poverty, and what is the individual contribution of each data source? Second, what are the characteristics of countries where models—including models trained using specific data sources—perform best? Third, do models trained within a country perform better than models trained on a pooled set of countries?

To answer these questions and improve previous literature models’ performance, we implement machine learning algorithms to estimate levels and changes of poverty across all countries with available Demographic and Health Survey (DHS) data, which spans Africa, Asia, the Americas, and Europe. DHS surveys are designed to be nationally representative, and as such, contain households across countries and across all levels of the income distribution. When estimating levels of poverty, we rely

on the most recent survey from 59 countries—spanning 63,854 survey clusters (survey clusters can be interpreted as villages or neighborhoods). In estimating changes in poverty, we rely on 33 countries that have at least two DHS survey rounds (we use the latest survey and the earliest since closest to 2000). DHS is not designed as a panel and different villages can be surveyed across different years. Following [48], we match clusters in one year to the nearest cluster in the survey in the previous year, only keeping pairs that are within 10km of each other. This process creates a synthetic panel of 7714 paired clusters. To train models, we rely on data that are globally and freely available to ensure that our approach can be easily replicated across countries. Models are trained on an asset-based wealth index. Following [48], we create a globally-comparable index by taking the first principle component of asset and household features when pooling all countries together (see appendix ?? for summary statistics of the global wealth index across countries). The asset-based wealth index provided by DHS is not comparable across countries, but our globally-comparable index has a high within-country correlation with the DHS-provided index across countries (see appendix ??).

Other papers have shed insight into the contribution of different data sources for poverty estimation. For example, [18] show that Facebook marketing data performs similarly to satellite imagery for poverty estimation in the Philippines, but worse in India where Facebook penetration is lower. [38] brings together a variety of data sources for poverty estimation—including CDR data, climate and environmental variables, and OpenStreetMap data, but focus on one country (Senegal). [8] rely on data across 56 countries using features from satellite imagery, Facebook connectivity data, and OpenStreet Map, and show that mobile connectivity data derived from Facebook is among the most predictive features of poverty compared to other features.

We train models on datasets that are (1) globally-available, (2) spatially-referenced, and (3) freely and publicly available, relying on seven sets of features. First, following much of the literature, we rely on daytime and nighttime satellite imagery. We use average values of nighttime lights and daytime imagery, and—following from [30]—use features from a transfer learning approach that uses a convolutional neural network (CNN) to train daytime imagery to predict nighttime lights. CNNs rely on the values and spatial configuration of values and can detect features such as roads and the spatial configuration of roads [49]. To take advantage of the multi-spectral nature of daytime imagery, we estimate three separate CNN models that use: (1) red, green and blue bands as inputs, (2) NDVI (a measure of vegetation) as an input, and (3) a built-up index as input. Second, we rely on synthetic aperture radar (SAR) data from Sentinel-1. SAR data is based on transmitting waves and measuring the strength and orientation of waves reflected back to the satellite sensor [21]. Different objects scatter waves differently, and research has use SAR data for uses such as vehicle detection, crop classification, and urban change monitoring [25, 1, 2]. Third, we rely on Facebook advertising data that measures

the proportion of the population on Facebook and the proportion of Facebook users across 34 different attributes (for example, the proportion that access Facebook using an expensive phone or the proportion that have an identified interest in luxury goods). Fourth, we use OpenStreetMap data to include the distance to and density of (1) points of interests (e.g., schools, health facilities, restaurants, etc) and (2) roads. Fifth, we use land cover data, including the proportion of land classified into 36 different land cover classes (e.g., built-up and cropland) from the European Space Agency GlobCover dataset, and—following from [38]—elevation and slope data. Sixth, as climate and weather contribute to food scarcity and agricultural production—particularly in rural areas—we rely on long-term climate features as well as annual temperature and precipitation [9]. Seventh, we rely on satellite-based measures of pollution. Pollution is included for two reasons. First, in some settings, poorer areas tend to be associated with higher levels of air pollution [23]. Second, pollution is an alternative source to nighttime lights to capture the extent of human activity [44].

Table 1 summarizes the data sources and features used to train the model. Not all data sources are sufficiently available across time. Consequently, we do not use data sources such as Facebook Marketing or OpenStreetMap data when training models that predict changes in poverty. In some cases we rely on similar but improved data sources for estimating levels of poverty; in particular, we rely on Sentinel-2 daytime imagery and VIIRS nighttime imagery to estimate CNN models when predicting levels of poverty, but rely on Landsat and DMSP-OLS to estimate CNN models when predicting changes in poverty.

We implement a supervised machine learning model to estimate the globally-comparable wealth index. We train models using the Extreme Gradient Boosting (XGBoost) algorithm [7]. XGBoost is a decision tree based algorithm that has proven to be one of the best performing machine learning algorithms across a variety of contexts. XGBoost implements multiple techniques to prevent against overfitting in high dimensional settings; this is an attractive property in our setting, where—when estimating levels of wealth—we rely on 443 features across the different data sources. We train models using all features and using only specific sets of features—such as models only trained on data from OpenStreetMap data, Facebook data, etc.

We use four different approaches for developing training and test sets: (1) within country models, training the model only using a sub-sample of country-wide data and testing the performance of the model using the remainder; (2) within continent estimation, where—for each country—we train a model using all other countries in the continent to estimate wealth in the country; (3) other continent estimation, where we train a model using each pair of continents and predict poverty in the other continent; and (4) global estimation, where we train a model on all other countries to predict poverty in a country. The approaches allow testing the trade-off between using less training data in a similar

context to more data in less similar contexts.

## 2 Results

### 2.1 Correlation of Features to Asset Wealth

Before turning to results from machine learning models, we first examine the correlation of features to levels and changes in wealth. Figure 1 shows the correlation of features to levels of wealth; panel A shows the distribution of within-country correlations for the feature with the highest median correlation in each dataset. Overall, individual features across most datasets see a relatively strong correlation with wealth; features from nighttime lights, OpenStreetMaps, land cover, all three CNN models, and Facebook marketing data all see median correlations near or above 0.5. Panels B-D show distributions of within-county correlations across features from Facebook, OpenStreetMaps and pollution datasets. Across all features from Facebook marketing data (panel B), the 25th percentile correlation is greater than zero; the proportion of Facebook users with interests in restaurants, luxury goods, and travel seeing the highest correlations of about 0.5 (appendix ?? shows the correlation of all Facebook features to the wealth index for each country). Panel C shows that the length of residential and all roads have the highest correlation with wealth; the 25th percentile of within-country correlations is above 0.5 for both variables. Among pollution variables, nitrogen dioxide is unique in having a large and positive correlation with wealth for most countries; however, other pollution variables show a much smaller correlation (while figure 1 shows the correlation of top features to wealth, see appendix ?? which shows the correlation between top features).

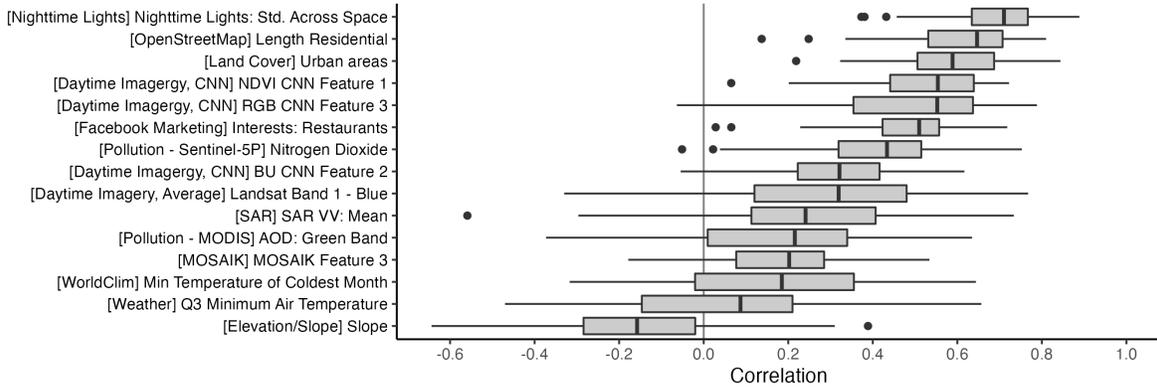
Figure 2 shows the correlation of changes in features to changes in wealth; the figure shows the distribution of within country correlations for the top two features with the highest median correlation in each dataset. Overall, correlations of changes are lower than that of correlations of levels. However, changes in nighttime lights, urban areas, and select features from CNNs generally see positive correlations for most countries, and with some countries seeing relatively high correlation coefficients. For example, the average correlation between changes in nighttime lights and the wealth index is 0.18, with a maximum correlation of 0.44 (in Mozambique).

Source	Time Span	Level/Change	Features for Poverty Estimation
<b>Daytime and Nighttime Satellite Imagery</b>			
VIIRS	2012- Present	Level	Nighttime lights: Average, standard deviation over time, and standard deviation over space
Harmonized DMSP-OLS and VIIRS	1992 - 2021	Changes	Nighttime Lights: Average and standard deviation over space
Landsat 7	1999 - 2021	Both	Spectral bands and indices (NDVI and build-up index): Average, standard deviation over time, and standard deviation over space
-	-	Changes	Convolutional neural network used to train daytime imagery on nighttime lights; features from CNN extracted
Sentinel-2	2015 - Present	Levels	Convolutional neural network used to train daytime imagery on nighttime lights; features from CNN extracted
<b>Synthetic Aperture Radar Data</b>			
Sentinel-1	2014 - Present	Levels	Synthetic aperture radar data, measuring average and standard deviation of VV and VH signals, and the ratio of the two—VV/VH. VV indicates vertical transmit, vertical receive, and VH indicates vertical transmit, horizontal receive.
<b>Facebook Marketing Data</b>			
Facebook	Present	Levels	Proportion of monthly active Facebook users according to select attributes (e.g., proportion of Facebook users with an iPhone)
<b>Roads and Points of Interest</b>			
OpenStreetMap	Present	Levels	(1) Number of points of interests (POIs) near survey (all and by type—e.g., restaurants, schools, health facilities, etc), (2) distance to nearest POI (all and by type), (3) Length of roads near survey (all and by type—e.g., trunk roads, primary roads, etc), and (4) distance to nearest road (all and by type)
<b>Land Cover and Type</b>			
ESA-GlobCover	1992-2018	Both	Proportion of area near survey classified according to 36 different land cover classes
Shuttle Radar Topography Mission (SRTM)	Time-Invariant	Levels	Average elevation and slope
<b>Weather and Climate</b>			
WorldClim	Average of 1970-2000	Levels	19 bioclimatic variables, including annual mean temperature, annual precipitation, mean temperature of wettest quarter, etc.
European Centre for Medium-Range Weather Forecasts: ERA5	1979-2020	Both	Average annual precipitation and temperature
<b>Pollution</b>			
Sentinel-5P	2018 - Present	Levels	Average pollution levels from six metrics: nitrogen dioxide, carbon monoxide,
MODIS	2000 - Present	Both	Aerosol optical depth

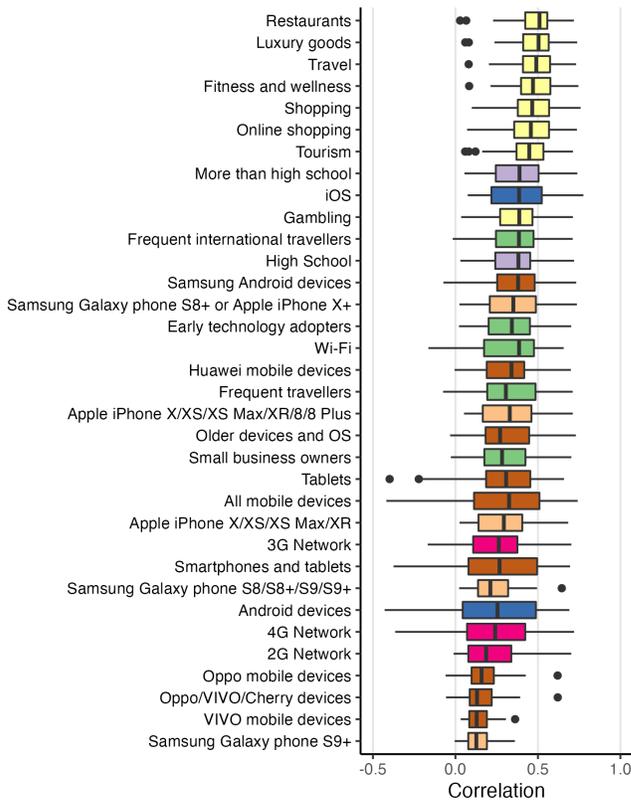
**Table 1:** Summary of data sources for poverty estimation

### A. Correlation of select variables to wealth index across countries

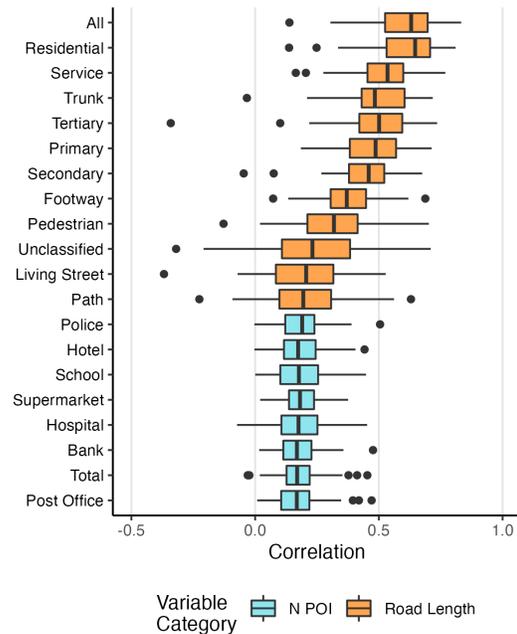
The variable with the highest median correlation for each dataset is shown



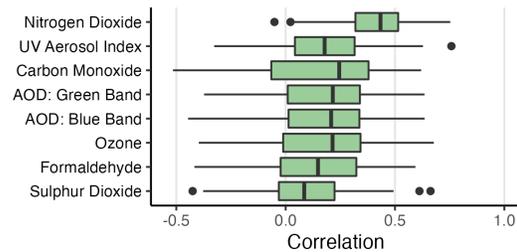
### B. Correlation of Facebook variables to wealth index



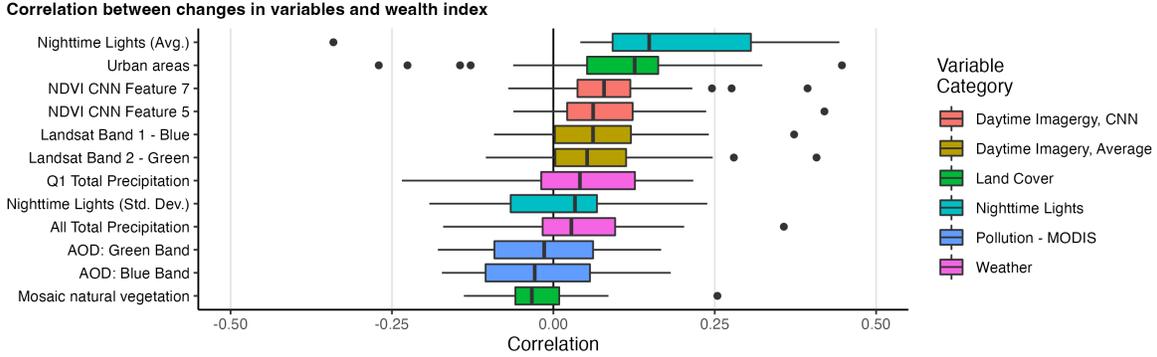
### C. Correlation of select OSM variables to wealth index



### D. Correlation of pollution variables to wealth index



**Figure 1:** Distribution of within-country correlations of select variables to levels of wealth. Correlations are computed at the cluster level using the latest survey year for each country. **Panel A** shows the distribution of within-country correlations of the feature with the highest median correlation across countries for each dataset. **Panel B** shows the correlation of all variables from the Facebook marketing data. **Panel C** shows select variables from OpenStreetMap data; the panel shows all variables of (1) the length of different classes of roads and (2) the number of different points of interests (POIs) near survey clusters. **Panel D** shows the correlation of all pollution variables from Sentinel-5P and MODIS; AOD variables are from MODIS and the other variables are from Sentinel-5P. The boxplots include: center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points beyond whiskers, outliers.



**Figure 2:** Distribution of within-country correlation of changes in select variables to changes in wealth, using clusters as the unit of analysis. We show the two variables with the highest median correlation for each dataset. The boxplots include: center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points beyond whiskers, outliers.

## 2.2 Machine Learning Model Performance

### 2.2.1 Predicting Levels of Wealth

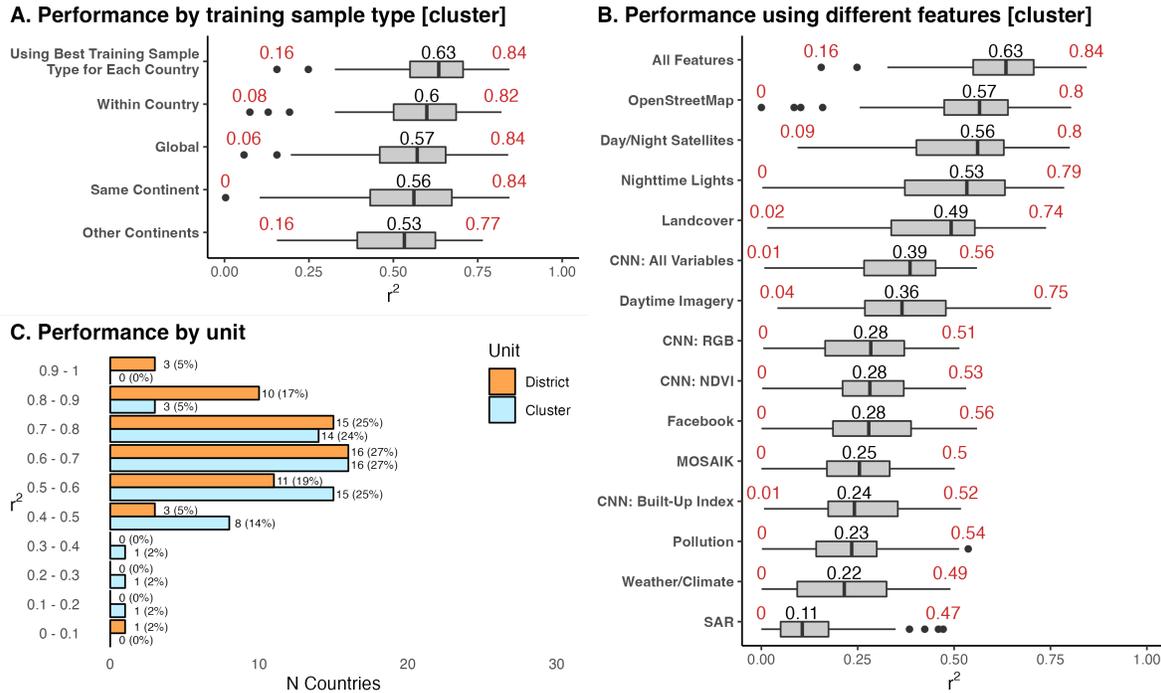
Models trained on only data within the country perform best, indicating that models work best when trained on more similar—albeit less—data (figure 3, panel A). Training on only within-country data performs best for 42% (N=25) countries, followed by training on countries in the same continent (35%; N = 21), followed by training on all other countries globally (15%; N = 9), followed by training on countries in other continents (7%; N = 4). Countries where within-country training performed best tend to have more clusters (median = 540; 75th percentile = 1128) than countries where training on other countries performed best (median = 399; 75th percentile = 551). This finding indicates that training on data from other countries is particularly advantageous for countries with less survey data to train models on.

Relying on models trained on all other countries and pooling data across countries, our model explains 61% of the variation in wealth at the cluster level and 75% of the variation when aggregating data to the district (second administrative division) level (figure 4, panels A and C; also see figure 3, panel C for comparison of cluster and district results). Separating across urban and rural clusters, the model explains 43% of the variation in poverty in urban areas and 50% of the variation in rural areas. The model performs best in Africa ( $r^2 = 0.68$ , on average across countries at the cluster level, and  $r^2 = 0.75$  at the district level), followed by the Americas ( $r^2 = 0.64$  at the cluster level;  $r^2 = 0.68$  at the district level) and Eurasia ( $r^2 = 0.54$  at the cluster level;  $r^2 = 0.6$  at the district level) (figure 4, panels B and D; appendix ?? shows scatterplots of true and estimated wealth for each country).

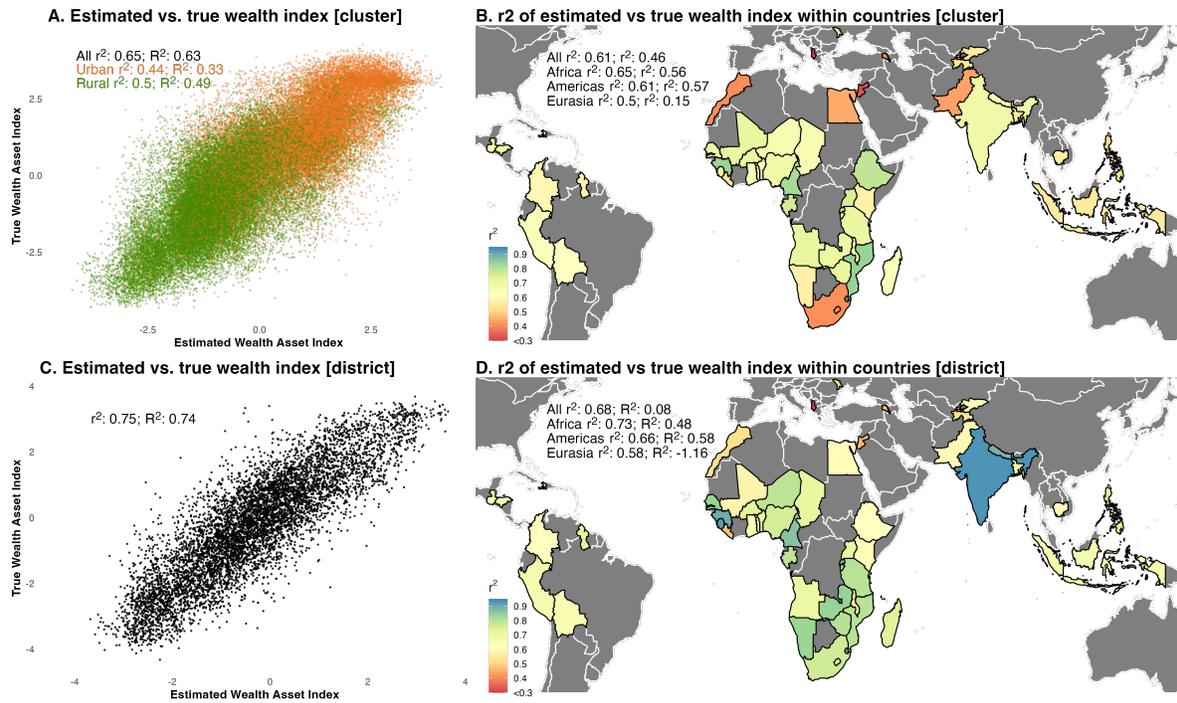
Relying on the best model for each country across all approaches—within country, other country, other continent, and global—models trained using all features see the best results, where the median  $r^2$  across countries is 0.67, and the minimum  $r^2$  is 0.25 (figure 3, panel B). Models using only features

from specific data sources—such as OpenStreetMaps or nighttime lights—only perform slightly worse, with a median  $r^2$  of 0.6; however, some countries see an  $r^2$  near zero when training on these individual data sources. While training models on only specific sets of features may work well for some countries, training models on features from multiple datasets helps to ensure good model performance across all countries (appendix ?? shows the importance of features from models trained on all features and shows similar results, which shows features from nighttime lights, land cover, and OpenStreetMaps as among the most important; appendix ?? shows model performance for each country using each set of features).

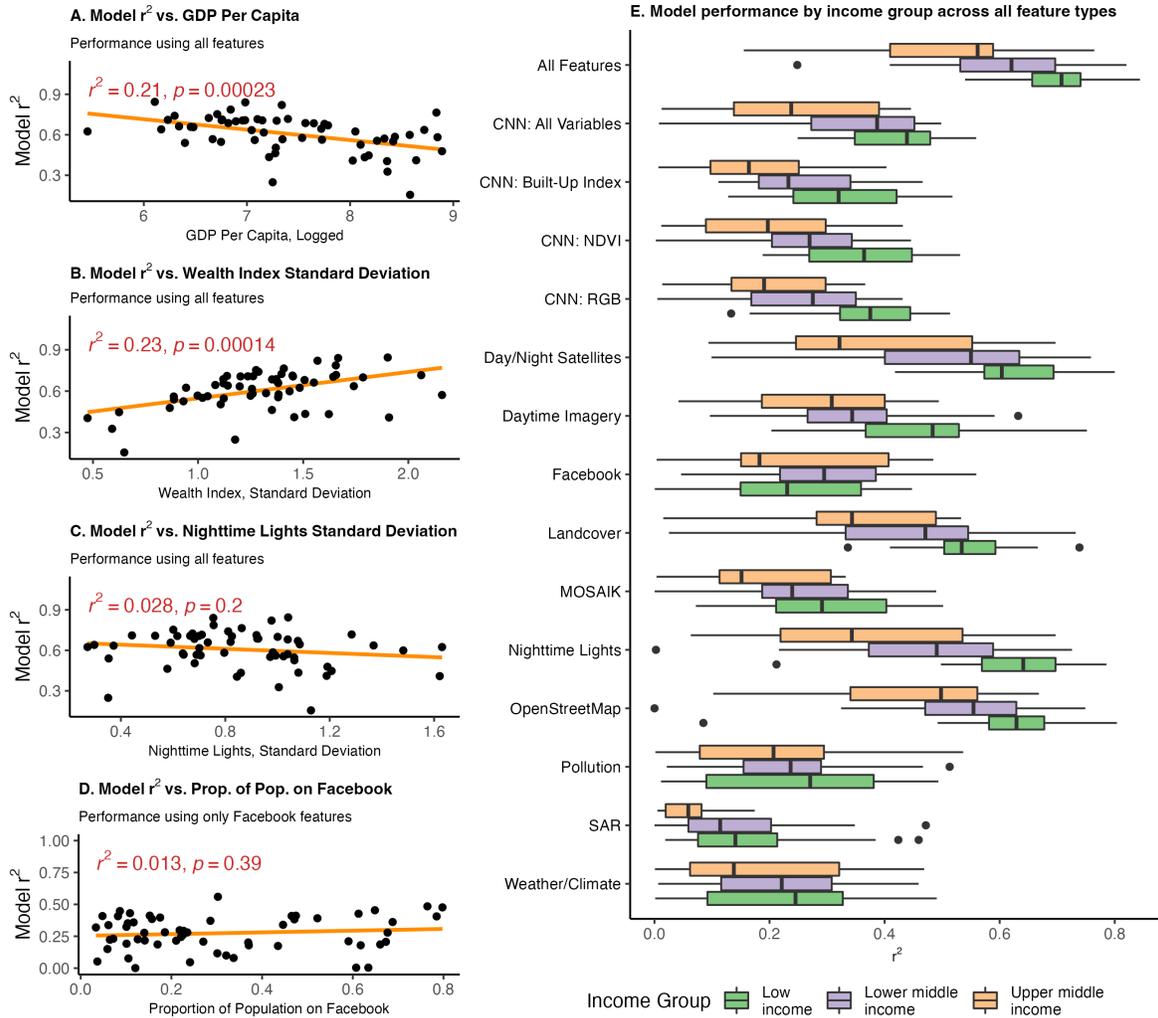
Figure 5 uses a number of country-level variables to explain where the models best explain wealth, relying on the most successful model for each country. Models work best in countries with more variation in the wealth index and lower GDP per capita (figure 5, panels A and B). Models trained only using Facebook marketing features perform similarly across countries with low and high proportions of the population on Facebook (figure 5, panels C and E). Models perform best in lower income countries across all feature types, with the exception of Facebook variables, pollution, and weather/climate variables—where models generally perform similarly across all income levels (figure 5, panel D).



**Figure 3:** Distribution of best model performance across countries, explaining levels of wealth. The black number shows the median and the red number shows the maximum  $r^2$ . **Panel A** shows model performance by the sample used to train countries. **Panel B** shows model performance when using different sets of features to train models. **Panel C** shows the distribution of model performance at the village and district level. Panels B and C use the model with the best performing training sample for each country (e.g., some countries may rely on models trained on just within-country data, while other countries may rely on models trained on all other countries). The boxplots include: center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points beyond whiskers, outliers.



**Figure 4:** Performance of best models predicting levels of wealth index. **Panel A** shows a scatterplot of the estimated and true wealth indices when pooling clusters across all countries, using models trained on all countries leaving out the country where wealth is being predicted. **Panel B** shows model performance when considering individual countries, using the best performing model for each country. **Panel C** and **Panel D** are similar to panels A and B, but using results aggregated at the district level.



**Figure 5:** Determinants of the variation in best model performance across countries, explaining levels of wealth. Models use clusters as the unit of analysis. **Panel A** shows the association between GDP per capita and model performance. **Panel B** shows the association between the wealth index standard deviation and model performance. **Panel C** shows the association between the nighttime lights standard deviation and model performance. **Panel D** shows the association between the proportion of the population on Facebook—measured using monthly active users divided by a country’s population—and model performance using only Facebook features to train the model. **Panel E** shows the distribution of model performance by income level using models trained across different feature sets. The boxplots include: center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points beyond whiskers, outliers.

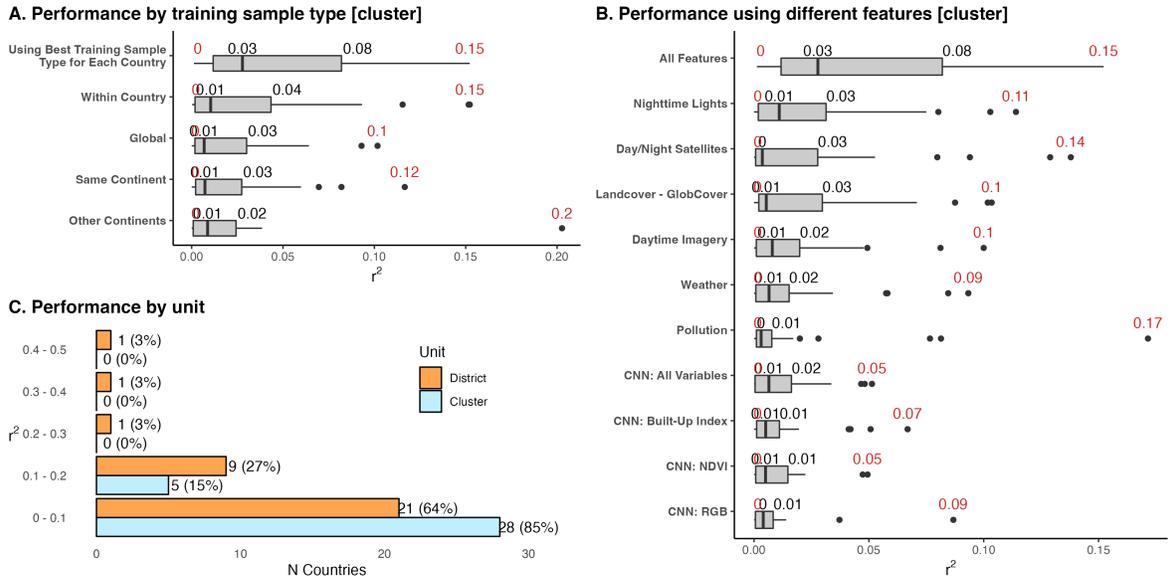
### 2.2.2 Predicting Changes in Wealth

Considering models using the best training sample for each country, the median model performance ( $r^2$ ) is 0.09 at the cluster level (the average  $r^2$  is 0.11; see figure 6, panel A). While this value is low, there is notable variation in model performance—in some countries, models explain up to 27% of the variation in changes in wealth. Models trained on all features perform best, followed by models trained on both day and nighttime imagery, followed by models trained on day nighttime imagery separately (figure 6, panel B; appendix ?? shows the importance of features from models trained on all features and shows similar results).

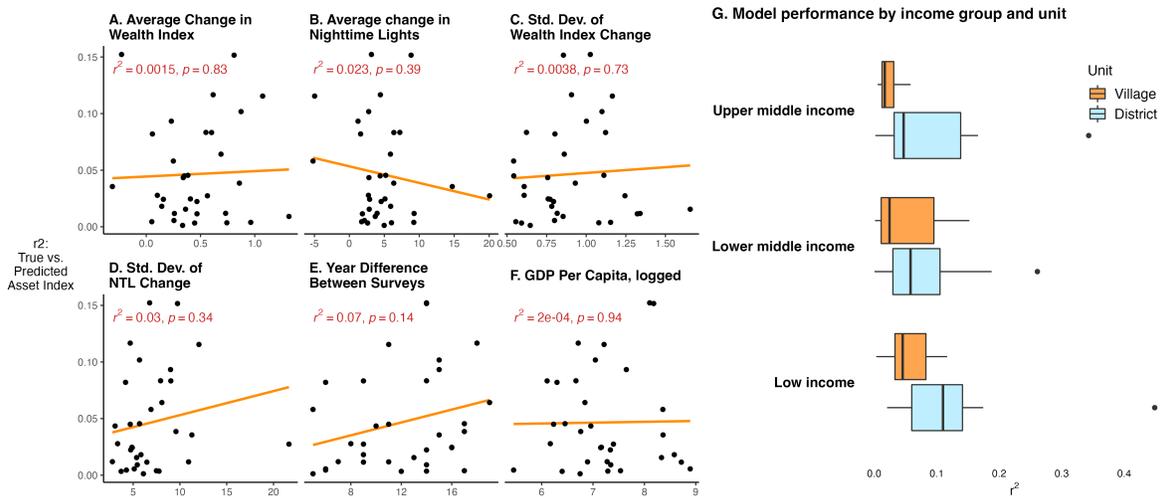
Model performance notably improves when aggregating data to the district (second administrative

unit) level (figure 6, panel C). On average across countries, the best model explains 10% of the variation in wealth; across low-income countries, the model explains 13% of the variation in wealth (figure 7). When pooling district-level data at the continent-level, models explain 29% of the variation in wealth in Eurasia, followed by 21% in Africa, followed by 10% in the Americas (figure 8; appendix ?? shows scatterplots of true and estimated changes in wealth for each country).

Figure 7 shows that countries that saw larger changes in wealth or nighttime lights (a common proxy for wealth) tended to see better model performance, although the association between changes in wealth and model performance is relatively weak (panels A and B). There is a stronger association between changes in the variation in wealth and nighttime lights and models performance—indicating that models work better in contexts where some locations in a country saw large increases in wealth and other locations did not (panels C and D). Countries where there is a larger year difference between surveys roughly see better model performance (panel E). In addition, lower income countries generally see better model performance (panels F and G).



**Figure 6:** Distribution of model performance across countries, explaining changes in wealth. The black number shows the median and the red number shows the maximum  $r^2$ . **Panel A** shows model performance by the sample used to train countries. **Panel B** shows model performance when using different sets of features to train models. **Panel C** shows the distribution of model performance at the village and district level. The boxplots include: center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points beyond whiskers, outliers.



**Figure 7:** Explaining variation in model performance across countries, explaining changes in wealth. The figure shows the association between model performance and average changes in wealth (**Panel A**), average changes in nighttime lights (**Panel B**), the standard deviation of the change in wealth (**Panel C**), standard deviation of the change in nighttime lights (**Panel D**), the years between surveys (**Panel E**), and current GDP per capita (**Panel F**). **Panel G** shows the distribution of model performance by income level, using villages and when aggregating to the district level. The boxplots include: center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points beyond whiskers, outliers.



**Figure 8:** Scatterplots of true and estimated changes in wealth

### 3 Discussion

This paper shows that globally, freely available, and spatially referenced data can accurately predict levels of wealth and—particularly at the district (second administrative level)—changes of wealth.

The latter is of stark importance when waves of data are separated by 5 years (on average), for lower income countries, or countries with subpopulations experiencing wealth shocks. The results show notable variation in model performance across countries, but show that they perform the best in countries where they can be most useful: countries with more variation—temporal and spatial—in wealth, and in lower income countries. The results also demonstrate the value added of using features from a variety of different data sources. The best performing models use data from all data sources. Models trained on features from just OpenStreetMaps, nighttime lights, or land cover data also perform well—on average—but there is large variation in model performance across countries, where the model explains virtually none of the variation in wealth in some countries when trained on these individual data sources. Models trained on features from all data sources guard against poor model performance in specific countries; across countries, the models estimating levels of wealth explain—at minimum—25% of the variation in wealth.

Some data sources provide features that are more interpretable than others. Beyond using features to explain changes in wealth, interpretable features can provide a richer context of characteristics of locations—this is particularly true for OpenStreetMaps and Facebook marketing data. For example, from OpenStreetMaps, the density of residential roads near a location is strongly correlated with wealth—but examining variation in levels and changes in residential roads could be useful to understand the level of amenities provided to richer households. Features from Facebook Marketing data can similarly provide useful understanding of populations—such as whether Facebook users are frequent travelers or whether they connect to Facebook from an expensive phone. When analyzing features individually, it is also important to consider potential biases. For example, OpenStreetMap relies on crowdsourced data such that some countries may have more complete data than others. Growth in roads could simply result from more complete data being added to OpenStreetMaps. Similarly, Facebook data could be biased as only certain segments of the population may be on Facebook (appendix ?? explores this bias by analyzing the association between a variable captured by both DHS and Facebook Marketing data—high school completion. We find that the variables move roughly together, particularly at the district level, which indicates that specific Facebook features can capture specific attributes of populations beyond wealth in some contexts). However, the strong association between OpenStreetMaps and Facebook variables with wealth shows that these data sources can capture on-the-ground dynamics—but additional work could inform further ground-truthing these data sources.

Overall, the paper shows the promise of leveraging new, globally available data sources for poverty estimation in contexts where it matters the most: for lower income countries, and when changes and levels of wealth vary within countries. All data sources come with their own sets of biases and

limitations. However, integrating data across different sources—from satellites to social network data to crowdsourced maps data from OpenStreetMaps—can help to overcome the limitations of any individual data source.

## 4 Methods

### 4.1 Preparation of survey data and wealth index

We rely on survey data from the Demographic and Health Surveys (DHS), which has become a common data source for training models for poverty estimation due to its wide coverage. We use data from all countries with available Standard DHS data with associated GPS coordinates of the survey clusters. For estimating levels of wealth, we rely on the most recent survey for each country, resulting in 63,854 survey clusters across 59 countries. For estimating changes in wealth, we rely on the most recent survey round and the oldest survey that was implemented closest to 2000, which results in a dataset capturing 7714 survey clusters across 33 countries. As the coverage of some data sources, particularly Landsat 7 and MODIS, started in 2000, we limit older surveys to those implemented near 2000 as the coverage of some data sources, particularly Landsat 7 and MODIS, started near 2000. This process results in including two countries with surveys implemented in 1998, so data captured in 2000 should still be indicative of 1998. DHS data are representative both nationally and at the the first administrative division level, where DHS provides a geographic coordinate for each cluster [39].

To protect privacy of respondents, DHS randomly displaces the true geographic location. Coordinates of urban clusters are displaced up to 2km, and most coordinates of rural clusters are displaced up to 5km—with 1% displaced up to 10km [5]. For most features, we extract values within 2.5 kilometers of the reported survey cluster location (e.g., average nighttime light values within 2.5 kilometers of the survey cluster). This process ensures that all urban clusters appear within the location used to extract values from features. Some rural clusters may appear outside the location used to extract values from features. However, especially in rural locations, we may expect positive spatial autocorrelation for a number of features—such that features are similar in nearby areas.

From DHS, we use a set of variables that capture the relative wealth of survey clusters. DHS provides a pre-computed wealth index, which is the first principle components of socio-economic attributes at the household level (e.g., roof material, owns assets such as a television, etc). The wealth index is computed within each country, so values from one country cannot be compared with an index from another country or within the same country over time. To remediate that issue, we follow an approach by [48] and use a set of socio-economic variables to construct a wealth index that is comparable across countries. The index is created by using the first principle component of the variables, a standard

approach to creating asset or wealth indices [45]. We use the following variables: ownership of assets including a television, fridge, motorbike, and car; access to electricity; quality of floor, wall and roof material (each categorized on a scale from 1-3 as natural, rudimentary or finished floor, roof or walls); time to get drinking water (categorized on a scale from 1-3 as more than 30 minutes, >0 to 30 minutes, and 0 minutes, or drinking water in household); whether the household has a flush toilet connected to the sewer system; number of people sleeping per bedroom (we divide the number of people listed in the household by the number of sleeping rooms and categorize this variable on a scale from 1-3 as more than 2 people/room, 1-2 people/room and less than 1 person/room); and the maximum number of years of education someone in the household has.

Not all the above variables are available for the earlier survey rounds. Consequently, when estimating levels of wealth—which rely on the most recent survey year—we include all the above variables when computing the principle component. When developing a wealth index to capture changes in wealth—which relies on earlier DHS surveys—we include a more limited set of features that are available in all survey rounds. Here, we use the ones above but exclude the number of people sleeping per bedroom, wall material, roof material, and time to get drinking water. Both indices (the index created from the full set of features and the index created from the limited set of features) are highly correlated with each other ( $\text{cor} = 0.981$ ). On average across countries, the index using the full set of features has a 0.917 correlation with the DHS wealth index, and the index using the limited set of features has a 0.864 correlation with the DHS wealth index.

## 4.2 Preparation of data for poverty estimation

We rely on a variety of sources of globally-available and spatially referenced data sources to use as features for estimating poverty. Datasets include raw satellite imagery, data derived from satellite imagery, private sector data and data derived from crowdsourcing. In this section we describe each of the datasets and variables extracted from the datasets; table 1 lists the different datasets and the features extracted from each dataset.

### 4.2.1 Nighttime Lights

Nighttime lights has been shown to be a strong proxy of local economic activity [12, 41, 10, 22, 27] and measures of welfare, including wealth estimates derived from DHS [46]. When estimating levels of wealth, we rely on the Visible Infrared Imaging Radiometer Suite (VIIRS), which has captured nighttime lights since 2012 at roughly a 500 meter resolution. We take the average value of nighttime lights within 2.5 kilometers of each survey cluster, the standard deviation of nighttime lights over space, and the standard deviation of nighttime lights over time (using monthly nighttime lights data)

using the year before, during, and after the survey year. When estimating changes in wealth, we rely on a dataset from [34] who harmonizes nighttime lights from DMSP-OLS (available from 1992 to 2013) and VIIRS (available from 2012 to present) into a single dataset. Raw DMSP-OLS data is at a lower resolution than VIIRS, and suffers from issues such as ‘blooming’ where light affects surrounding pixels [14]. The harmonized dataset uses VIIRS data to simulate DMSP-OLS-like data, so that trends in nighttime lights can be measured across DMSP-OLS and VIIRS time periods. The harmonized dataset is available annually from 1992 to 2021; we take the average and standard deviation values of nighttime lights within 2.5 kilometers of each survey cluster.

#### 4.2.2 CNN Features from Daytime Imagery Predicting Nighttime Lights

We follow previous research that uses convolutional neural networks (CNNs) to extract features from satellite imagery. CNNs capture the spatial configuration of images, not just average values from imagery; consequently, CNNs may capture features such as roads, buildings, vegetation and how these are arranged (e.g., buildings in neat rows or a more disorganized layout) [49]. We follow previous studies that use a transfer learning approach, whereby a CNN model using daytime imagery as features is used to predict the magnitude of nighttime lights [30, 47]. The features created from this model are then used in a separate model to estimate poverty. Given that nighttime lights is widely recognized as a strong proxy for economic activity and wealth ([11]), a CNN predicting nighttime lights will develop features that are also relevant for capturing poverty.

For estimating levels of wealth, we rely on daytime imagery from Sentinel-2 [13]. Sentinel-2 has a 10 meter resolution; surface reflectance data is made available since 2017 and captures imagery roughly every two weeks. For each 10 meter pixel, we use the median value across images from 2017-2020. Using data across multiple years follows from [48], who note that using a longer time period minimizes the impact of cloud cover and minimizes distortions by seasonal or short-run variations in imagery values. One disadvantage of using Sentinel data is that many DHS surveys were administered before 2017. However, previous research faces a similar issue and observes strong results, indicating that imagery from later years still provides useful information about poverty in past years. For example, [30] relies on satellite imagery from Google Maps from 2013 to 2015 to estimate poverty in 2010-2013. An alternative approach would be to use Landsat data, where imagery goes back decades; however, we opt for Sentinel as it has a finer spatial resolution than Landsat’s 30 meter resolution. For estimating changes in wealth, we do rely on Landsat imagery. We rely on Landsat 7, which provides imagery from 1999 to 2022 at a 30 meter resolution. Similar to how we process Sentinel-2 data, we use data from the year before, during, and after the survey year—taking the median value for each pixel across these years.

We use daytime imagery to predict nighttime lights. When estimating levels of wealth, we rely on VIIRS; when estimating changes, we rely on the DMSP-OLS harmonized dataset from [34]. We take average nighttime lights over the same time period as daytime imagery is taken. We then group nighttime lights radiance into five equally sized groups, representing low to high nighttime lights. Figure 9 shows example daytime images from Sentinel across these five groups. For the CNN, we use the pre-trained VGG16 model and further train the model to predict nighttime lights.

For both Sentinel and Landsat data, we use a tiles of 224 x 224 pixels, which is the tile size used to train the VGG16 model. Consequently, for estimating levels of nighttime lights, we use a daytime grid that is 2.4 x 2.4 kilometers (224 \* 10 meter resolution), and we use average nighttime lights within 2.4 kilometers of the survey cluster. For estimating changes of nighttime lights, we use a daytime grid that is 6.72 kilometers by 6.72 kilometers (224 \* 30 meter resolution), and we use average nighttime lights within 6.72 kilometers of the survey cluster.

We train three separate CNN models using different sets of daytime imagery. First, following typical approaches, we train a model using the visible red, green and blue bands. For the second and third models we take advantage of Landsat and Sentinel capturing multiple spectral bands beyond those visible to the human eye. We compute two indices using multiple bands: the normalized difference vegetation index (NDVI), a common metric for capturing biomass or vegetation, and the built up index (BU), a metric to capture human settlement or built-up areas [26]. NDVI and BU are computed as:

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (1)$$

$$BU = \frac{SWIR - NIR}{SWIR + NIR} - NDVI \quad (2)$$

where NIR is the near-infrared band and SWIR is the shortwave infrared band. Our second CNN model uses NDVI as an input and our third model uses BU as an input. The pre-trained VGG16 model requires three layers of inputs; consequently, we repeat the NDVI and BU bands into three layers in order to still take advantage of the pre-trained model

**Daytime Images from Areas with Low Nighttime Lights**



**Daytime Images from Areas with Medium Low Nighttime Lights**



**Daytime Images from Areas with Medium Nighttime Lights**



**Daytime Images from Areas with Medium High Nighttime Lights**



**Daytime Images from Areas with High Nighttime Lights**



**Figure 9:** Example daytime images from different nighttime lights groups

### 4.2.3 Daytime Imagery

In addition to relying on CNN models to extract features from daytime imagery, we also test a more simple use of daytime imagery: extracting average and the standard deviation of values from imagery. We take average values of spectral bands within 2.5 kilometer buffers. We rely on Landsat 7 when estimating levels and changes. Landsat 7 captures daytime imagery across 6 different spectral bands (e.g., red, green, blue and near-infrared imagery), where data from most spectral bands are captured at a 30 meter resolution. We take the average and standard deviation of values from each spectral band near each survey location; as with nighttime lights, we take the average using imagery from the year before, during and after the survey was implemented. In addition to using the raw values from the spectral bands, we also use average values of NDVI and BU. NDVI and BU are more interpretable metrics than the individual bands and have a more expected relation with wealth. For example, more built-up (urban) area may see greater wealth than rural areas. However, while the individual bands do not provide as interpretable of metrics, the bands and combinations can capture factors related to poverty.

### 4.2.4 Facebook Advertising Data

[19] demonstrate the use of Facebook advertising data for poverty estimation using data from the Philippines and India. They use the Facebook Marketing API to query the number of Facebook monthly active users that match certain criteria, such as the number using certain devices (iOS vs. Android), use of high-end devices (latest iPhone or Samsung Galaxy Phones), among other characteristics.

Following [19], we query the number of monthly active Facebook users (MAUs) across a variety of attributes around each survey locations. We query 34 different attributes, from categories including reported interests (e.g., interest in online shopping), behaviors (e.g., frequent travellers), how one accesses Facebook (e.g., through an expensive phone like a new iPhone), among other categories. We compute the proportion of Facebook users across all attributes. The Facebook Marketing API allows querying areas with a minimum 1km radius, and to protect privacy does not provide MAU values less than 1000; consequently, areas with low Facebook usage are reported to have an MAU of 1000.

Choosing the radius to query Facebook data requires balancing the radius being large enough to have sufficient Facebook users (i.e., over 1000) but not extending too far beyond the survey cluster location. [19] use a 2km radius for urban areas and a 5km radius for rural survey locations in Philippines and—to account for lower Facebook penetration in India—use larger radii of 5km and 10km for urban and rural locations in India. We use an alternative, data driven approach to define the radius for querying data both because: (1) a goal of poverty estimation is to extrapolate to areas beyond survey locations, and we may not necessarily know whether a location should be classified as urban or rural

and (2) we query data from a larger set of countries which would make defining custom radii for each country cumbersome.

To choose the radius to query Facebook data, for each survey location, we check MAU of all Facebook users using a radius of 2km, 5km and 10km. We use the smallest radius that has at least 2000 MAUs (if none have 2000 MAUs, we use the 10km radius). We use a threshold of 2000 MAUs as we are interested in the proportion of Facebook users across characteristics; using a threshold above 1000 for all Facebook users helps to ensure that MAUs of Facebook users across select characteristics also have a value above 1000. 2km, 5km and 10km radii follow from the random displacement of DHS location displacement values; urban clusters are displaced up to 2km, and most rural clusters are displaced up to 5km—with 1% displaced up to 10km. Our strategy for choosing radii to query Facebook for each cluster allows urban clusters to use a 5km or 10km if there is low Facebook penetration. In addition, our strategy also allows a rural cluster to use a 2km radius; here, there is a chance that the radius chosen for Facebook does not cover the true DHS cluster location. However, given that there is likely some degree of positive spatial autocorrelation of poverty levels, if the queried location does not capture the true cluster location, the queried values likely still provide values that would be roughly similar to those at the true cluster location.

#### 4.2.5 OpenStreetMaps

OpenStreetMap (OSM) is a crowdsourced initiative to create a global geographical database [24]. OSM contains data on roads, buildings, points of interests (e.g., schools, health facilities, etc) and other features such as bodies of water. Existing research has incorporated measures from OSM into poverty estimation models as features such as accessibility to different road types and points of interest can be associated with socioeconomic levels [38, 50, 35, 33]. For example, more wealthy areas may be both situated closer to services such as health clinics and schools and have better access to services through a large road network. In addition, the number of points of interest and the size of the road network near survey locations can also help indicate the sparsity or density of the location.

We adapt strategies for extracting features from OSM data from existing work, particularly [38]. We first extract the number of each type of point of interest near each survey location (within 2.5km) and the minimum distance to each point of interest. OSM contains hundreds of types of points of interest; we use the 50 most common points of interest, determined by summing the number of points of interest near each survey location across all survey locations. Second, we extract the length of each road type near each survey location and compute the minimum distance to each road type. Road types include, motorways, trunk, primary, secondary, tertiary, service, pedestrian, among others.

We only use OSM data to measure levels of poverty. Historical OSM data can be downloaded,

so OSM could potentially also be used to explain changes in wealth. However, changes in OSM data could be due to both (1) new roads or points of interest or (2) OSM data being updated to better reflected currently existing roads or points of interest. More recent changes in OSM may more reflect new roads or points of interest. However, as OSM was founded in 2006 and many of the first round DHS surveys in our dataset come from the early 2000s, changes from the early 2000s to near the present may more reflect new data simply being entered into OSM. Consequently, we only use OSM data when estimating levels of wealth.

#### **4.2.6 Land Cover and Characteristics**

Land cover and characteristics provide useful context for understanding socioeconomic characteristics. For example, land cover can indicate the extent of urban and cropland areas. We use the Globcover land cover dataset, which is produced by the European Spatial Agency and is available annually from 1992 to 2018 at 300m resolution [16, 17]. It classifies each pixel into one of the 36 land cover classes defined with the United Nations Land Cover Classification System (e.g, build-up/urban, cropland, etc). We extract the proportion of land classified according to each class near each survey cluster. In addition, following [38], we include physical land characteristics including elevation and slope. We use slope and elevation data from the NASA Shuttle Radar Topography Mission (SRTM), which provides imagery at a 30m resolution.

#### **4.2.7 Climate and Weather**

Climate and weather contribute to food scarcity and agricultural production, particularly in rural areas. Following from [38], we include long-term climate features from the WorldClim version 2 dataset, which captures annual trends, seasonality, and extreme values of temperature and precipitation at a roughly 1km resolution; values are averaged from data from 1970 to 2000 [29]. In addition to long-term climate trends, we include temperature and precipitation features captured the year the survey was implemented from the Copernicus Climate Change Service [42].

#### **4.2.8 Pollution**

Previous research has shown that in many geographic regions, poorer areas tend to be associated with higher levels of air pollutants [23]. Following this literature, we use measures of pollution around survey locations from two data sources. First, we use pollution measures from Sentinel-5 Precursor (Sentinel-5P), a satellite designed specifically to monitor the Earth’s atmosphere launched in 2017 [43]. From Sentinel 5-P, we use pollution metrics including a UV aerosol index, Carbon Monoxide, Formaldehyde, Ozone, Sulphur Dioxide, and Methane. Together, these sources capture pollution generating activities

such as fossil fuel and biomass burning, and traffic and industrial activity. We use average pollution measures captured from 2018-2020. To also capture pollution levels during the year of the survey, as our second source we use a measure of aerosol optical depth (AOD) from the Moderate Resolution Imaging Spectroradiometer (MODIS), which provides imagery since 2000 [32]. AOD is a measure of the amount of aerosols in the atmosphere, and has been shown to correlate strongly with measures of particulate matter [20].

#### 4.2.9 Synthetic Aperture Radar Data

We use synthetic aperture radar (SAR) data from Sentinel-1. SAR data is based on transmitting waves and measuring the strength and orientation of waves reflected back to the satellite sensor [21]. As SAR sensors both transmit and receive signals, SAR data is unaffected by cloud cover—unlike daytime and nighttime satellite imagery which receives optical data reflected by the earth’s surface. Different objects scatter waves differently. For example, buildings and metallic objects produce a strong signal back to the sensor, and flat surfaces produce a weak signal [25]. Research has used SAR data for uses such as vehicle detection, crop classification, and urban change monitoring [25, 1, 2, 31, 4].

We use average and standard deviation of VV and VH bands, and compute the average and standard deviation of VV/VH. VV indicates waves that are transmitted vertically and received vertically, and VH indicates waves that are transmitted vertically and received horizontally. We use a 2.5km buffer to compute values. Using all three metrics—VV, VH, and the ratio of the two—helps to further distinguish objects or land types. For example, vertical objects near smooth surfaces will give a strong VV signal but a weak VH signal [15].

## 5 Machine learning models

We train machine learning models using the Extreme Gradient Boosting algorithm (XGBoost) [7]. XGBoost comes with a number of parameters that can be optimized. We use a grid search to choose the model with best performance across four parameters: maximum depth of the tree (2, 5, 6 and 10), learning rate (0.3), the number of boosting rounds (50, 100, and 500), and the proportion of the training data to randomly sample prior to growing trees (0.3, 0.6, and 1),

We test four approaches for defining train and test sets:

- **Within Country Estimation:** We train models only using data within each country. We divide the country into five folds, training the model on four folds and predicting poverty on the remaining fold. Folds are created by randomly assigning each second administrative division to one of the five folds; the survey cluster is assigned the fold by using the administrative division it

falls in. We use the Database of Global Administrative Areas (GADM) for boundaries of administrative divisions. In cases where the second administrative level isn't available, we rely on the first administrative unit; this occurred in four cases (Armenia, Comoros, Lesotho and Moldova). In one case (Comoros) we use three folds as there were only three second administrative divisions. We create folds using administrative divisions to mitigate against nearby survey clusters being both in the train and test sets.

- **Within Continent Estimation:** For each country, we train a model using all other countries in the same continent to predict poverty in the country. Given that some continents only have few countries (Europe has two countries with data available from DHS), we use the following continent groupings: (1) Americas (6 countries); (2) Euroasia (2 countries in Europe, 12 countries in Asia and 1 country in Oceania); and (3) Africa (32 countries).
- **Other Continents Estimation:** We train a model using each pair of continents and predict poverty in the other continent.
- **Global Estimation:** For each country, we train a model on all other countries to predict poverty in the country.

The approaches allow testing the balance between using data in a similar context to extending training data to using more data but in less similar contexts. At one extreme, we rely only on data within each country to train a model for poverty estimation; this approach uses the least amount of data, but helps ensure the training set comes from a similar context as the test set (while still ensuring that training and test sets do not spatially overlap). At the other extreme is leveraging data on all countries to predict poverty in a single country; while this takes advantage of using more data, it may suffer from introducing training data that may be less helpful for predicting poverty in a specific context (e.g., training with data on Armenia may not be helpful for predicting poverty in Lesotho). Training with data from other other countries within the same continent provides a middle ground approach. Training with data from countries in different continents directly tests whether training with data on less similar countries can still provide useful poverty estimates.

## 6 Data availability

Data to replicate all findings in the paper are available at

<https://github.com/dime-worldbank/big-data-poverty-estimation>

## 7 Code availability

Code to replicate all findings in the paper are available at

<https://github.com/dime-worldbank/big-data-poverty-estimation>

## References

- [1] Sebastián Amherdt, Néstor Cristian Di Leo, Sebastián Balbarani, Ayelen Pereira, Cecilia Cornero, and María Cristina Pacino. Exploiting sentinel-1 data time-series for crop classification and harvest date detection. *International Journal of Remote Sensing*, 42(19):7313–7331, 2021.
- [2] Timo Balz, Prosper Washaya, and Michael Jendryke. Urban change monitoring using globally available sentinel-1 imagery. In *2018 International Workshop on Big Geospatial Data and Data Science (BGDDS)*, pages 1–4, 2018.
- [3] World Bank. Poverty and inequality platform (version 20220909\_2017\_01\_02\_prod). Technical report, World Bank, 2022. World Bank Group. [www.pip.worldbank.org](http://www.pip.worldbank.org). Accessed Jan 2023.
- [4] Bernhard Bauer-Marschallinger, Senmao Cao, Claudio Navacchi, Vahid Freeman, Felix Reuß, Dirk Geudtner, Björn Rommen, Francisco Ceba Vega, Paul Snoeij, Evert Attema, Christoph Reimer, and Wolfgang Wagner. The normalised sentinel-1 global backscatter model, mapping earth’s land surface with c-band microwaves. *Scientific Data*, 8(1):277, Oct 2021.
- [5] Clara R Burgert, Josh Colston, Thea Roy, and Blake Zachary. *Geographic displacement procedure and georeferenced data release policy for the Demographic and Health Surveys*. ICF International, 2013.
- [6] Marshall Burke, Anne Driscoll, David B. Lobell, and Stefano Ermon. Using satellite imagery to understand and promote sustainable development. *Science*, 371(6535):eabe8628, 2021.
- [7] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’16*, page 785–794, New York, NY, USA, 2016. Association for Computing Machinery.
- [8] Guanghua Chi, Han Fang, Sourav Chatterjee, and Joshua E Blumenstock. Microestimates of wealth for all low-and middle-income countries. *Proceedings of the National Academy of Sciences*, 119(3):e2113658119, 2022.
- [9] Shouro Dasgupta and Elizabeth J. Z. Robinson. Attributing changes in food insecurity to a changing climate. *Scientific Reports*, 12(1):4709, Mar 2022.

- [10] Christopher N.H. Doll, Jan-Peter Muller, and Jeremy G. Morley. Mapping regional economic activity from night-time light satellite imagery. *Ecological Economics*, 57(1):75–92, 2006.
- [11] Dave Donaldson and Adam Storeygard. The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives*, 30(4):171–98, November 2016.
- [12] Dave Donaldson and Adam Storeygard. The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives*, 30(4):171–98, November 2016.
- [13] Matthias Drusch, Umberto Del Bello, Sébastien Carlier, Olivier Colin, Veronica Fernandez, Ferran Gascon, Bianca Hoersch, Claudia Isola, Paolo Laberinti, Philippe Martimort, et al. Sentinel-2: Esa’s optical high-resolution mission for gmes operational services. *Remote sensing of Environment*, 120:25–36, 2012.
- [14] Christopher D. Elvidge, Kimberly Baugh, Mikhail Zhizhin, and Feng Chi Hsu. Why VIIRS data are superior to DMSP for mapping nighttime lights. In *Proceedings of the Asia-Pacific Advanced Network*, 2013.
- [15] ESRI. Interpretation of sar data for flood mapping.
- [16] European Space Agency. 300 m annual global land cover time series from 1992 to 2015, 2017.
- [17] European Space Agency. New release of the c3s global land cover products for 2016, 2017 and 2018 consistent with the cci 1992 – 2015 map series, 2019.
- [18] Masoomali Fatehkia, Benjamin Coles, Ferda Ofli, and Ingmar Weber. The relative value of facebook advertising data for poverty mapping. *Proceedings of the International AAAI Conference on Web and Social Media*, 14(1):934–938, May 2020.
- [19] Masoomali Fatehkia, Isabelle Tingzon, Ardie Orden, Stephanie Sy, Vedran Sekara, Manuel Garcia-Herranz, and Ingmar Weber. Mapping socioeconomic indicators using social media advertising data. *EPJ Data Science*, 9(1):22, Jul 2020.
- [20] Mikalai Filonchyk, Haowen Yan, Zhongrong Zhang, Shuwen Yang, Wei Li, and Yanming Li. Combined use of satellite and surface observations to study aerosol optical depth in different regions of china. *Scientific Reports*, 9(1):6174, Apr 2019.
- [21] Dirk Geudtner, Ramón Torres, Paul Snoeij, Malcolm Davidson, and Björn Rommen. Sentinel-1 system capabilities and applications. In *2014 IEEE Geoscience and Remote Sensing Symposium*, pages 1457–1460, 2014.

- [22] Tilottama Ghosh, Sharolyn J. Anderson, Christopher D. Elvidge, and Paul C. Sutton. Using nighttime satellite imagery as a proxy measure of human well-being. *Sustainability*, 5(12):4988–5019, 2013.
- [23] Anjum Hajat, Charlene Hsia, and Marie S. O’Neill. Socioeconomic disparities and air pollution exposure: a global review. *Current Environmental Health Reports*, 2(4):440–450, Dec 2015.
- [24] Mordechai Haklay and Patrick Weber. Openstreetmap: User-generated street maps. *IEEE Pervasive Computing*, 7(4):12–18, 2008.
- [25] Michael Harner, Austen Groener, and Mark Pritt. Detecting the presence of vehicles and equipment in sar imagery using image texture features. In *2019 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, pages 1–6, 2019.
- [26] Chunyang He, Peijun Shi, Dingyong Xie, and Yuanyuan Zhao. Improving the normalized difference built-up index to map urban built-up areas using a semiautomatic segmentation approach. *Remote Sensing Letters*, 1(4):213–221, 2010.
- [27] J. Vernon Henderson, Adam Storeygard, and David N. Weil. Measuring economic growth from outer space. *American Economic Review*, 102(2):994–1028, April 2012.
- [28] Jonathan Hersh, Ryan Engstrom, and Michael Mann. Open data for algorithms: mapping poverty in belize using open satellite derived features and machine learning. *Information Technology for Development*, 27(2):263–292, 2021.
- [29] Robert J. Hijmans, Susan E. Cameron, Juan L. Parra, Peter G. Jones, and Andy Jarvis. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15):1965–1978, 2005.
- [30] Neal Jean, Marshall Burke, Michael Xie, W Matthew Davis, David B Lobell, and Stefano Ermon. Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301):790–794, 2016.
- [31] Deepak Kumar. Urban objects detection from c-band synthetic aperture radar (sar) satellite images through simulating filter properties. *Scientific Reports*, 11(1):6241, Mar 2021.
- [32] HJ Lee, Y Liu, BA Coull, J Schwartz, and P Koutrakis. A novel calibration approach of modis aod data to predict pm 2.5 concentrations. *Atmospheric Chemistry and Physics*, 11(15):7991–8002, 2011.
- [33] Kamwoo Lee and Jeanine Braithwaite. High-resolution poverty maps in sub-saharan africa, 2021.

- [34] Xuecao Li, Yuyu Zhou, Min Zhao, and Xia Zhao. A harmonized global nighttime light dataset 1992–2018. *Sci Data*, 7(168), 2020.
- [35] Nyan Lin Htet, Waree Kongprawechnon, Suttipong Thajchayapong, and Tsuyoshi Isshiki. Machine learning approach with multiple open-source data for mapping and prediction of poverty in myanmar. In *2021 18th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, pages 1041–1045, 2021.
- [36] Ian McCallum, Christopher Conrad Maximillian Kyba, Juan Carlos Laso Bayas, Elena Moltchanova, Matt Cooper, Jesus Crespo Cuaresma, Shonali Pachauri, Linda See, Olga Danylo, Inian Moorthy, et al. Estimating global economic well-being with unlit settlements. *Nature communications*, 13(1):1–8, 2022.
- [37] Christopher Njuguna and Patrick McSharry. Constructing spatiotemporal poverty indices from big data. *Journal of Business Research*, 70:318–327, 2017.
- [38] Neeti Pokhriyal and Damien Christophe Jacques. Combining disparate data sources for improved poverty prediction and mapping. *Proceedings of the National Academy of Sciences*, 114(46):E9783–E9792, 2017.
- [39] The DHS Program. Dhs methodology. Accessed January 20, 2023.
- [40] Jessica E. Steele, Pål Roe Sundsøy, Carla Pezzulo, Victor A. Alegana, Tomas J. Bird, Joshua Blumenstock, Johannes Bjelland, Kenth Engø-Monsen, Yves-Alexandre de Montjoye, Asif M. Iqbal, Khandakar N. Hadiuzzaman, Xin Lu, Erik Wetter, Andrew J. Tatem, and Linus Bengtsson. Mapping poverty using mobile phone and satellite data. *Journal of The Royal Society Interface*, 14(127):20160690, 2017.
- [41] Paul C. Sutton and Robert Costanza. Global estimates of market and non-market values derived from nighttime satellite imagery, land cover, and ecosystem service valuation. *Ecological Economics*, 41(3):509–527, 2002.
- [42] Jean-Noël Thépaut, Dick Dee, Richard Engelen, and Bernard Pinty. The copernicus programme and its climate change service. In *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 1591–1593, 2018.
- [43] JP Veefkind, I Aben, K McMullan, H Förster, J De Vries, G Otter, J Claas, HJ Eskes, JF De Haan, Q Kleipool, et al. Tropomi on the esa sentinel-5 precursor: A gmes mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote sensing of environment*, 120:70–83, 2012.

- [44] Zander S. Venter, Kristin Aunan, Sourangsu Chowdhury, and Jos Lelieveld. Covid-19 lockdowns cause global air pollution declines. *Proceedings of the National Academy of Sciences*, 117(32):18984–18990, 2020.
- [45] Seema Vyas and Lilani Kumaranayake. Constructing socio-economic status indices: how to use principal components analysis. *Health Policy and Planning*, 21(6):459–468, 10 2006.
- [46] Nils B Weidmann and Sebastian Schutte. Using night light emissions for the prediction of local wealth. *Journal of Peace Research*, 54(2):125–140, 2017.
- [47] Michael Xie, Neal Jean, Marshall Burke, David Lobell, and Stefano Ermon. Transfer learning from deep features for remote sensing and poverty mapping. In *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.
- [48] Christopher Yeh, Anthony Perez, Anne Driscoll, George Azzari, Zhongyi Tang, David Lobell, Stefano Ermon, and Marshall Burke. Using publicly available satellite imagery and deep learning to understand economic well-being in africa. *Nature Communications*, 11(1):2583, May 2020.
- [49] Qianqian Zhang, Qingling Kong, Chao Zhang, Shucheng You, Hai Wei, Ruizhi Sun, and Li Li. A new road extraction method using sentinel-1 sar images based on the deep fully convolutional neural network. *European Journal of Remote Sensing*, 52(1):572–582, 2019.
- [50] Xizhi Zhao, Bailang Yu, Yan Liu, Zuoqi Chen, Qiaoxuan Li, Congxiao Wang, and Jianping Wu. Estimation of poverty using random forest regression with multi-source data: A case study in bangladesh. *Remote Sensing*, 11(4), 2019.