Population, Space and Place

WILEY

RESEARCH ARTICLE OPEN ACCESS

How Accurate Are High Resolution Settlement Maps at Predicting Population Counts in Data Scarce Settings?

Edith Darin[1,2] | Ridhi Kashyap[3,4] | Douglas R. Leasure[1,4]

[1]Nuffield Department of Population Health, University of Oxford, Oxford, UK | [2]St Catherine's College, University of Oxford, Oxford, UK | [3]Department of Sociology, University of Oxford, Oxford, UK | [4]Nuffield College, University of Oxford, Oxford, UK

Correspondence: Edith Darin (edith.darin@demography.ox.ac.uk)

Received: 24 July 2024 | **Revised:** 3 July 2025 | **Accepted:** 25 July 2025

Funding: This study was supported by the Leverhulme Trust under Grant RC-2018-003 for the Leverhulme Centre for Demographic Science and the ESRC Impact Acceleration Account under Grant 2209-KEA-835.

Keywords: Bayesian | census-independent estimation | data-scarce context | population | random forest | satellite imagery | settlement map

ABSTRACT

Despite the recent milestone of the world population surpassing 8 billion, disparities in population data reliability persist, with many countries facing outdated or incomplete census data. Such inaccuracies have far-reaching implications for various sectors, including public health, urban planning, and resource allocation. The study leverages the rich data environment provided by the detailed 2018 Colombian census data and its coverage indicator, which create a high-quality controlled environment to assess the performance of census-independent population estimation approaches. Drawing from a diverse range of environmental landscapes in Colombia, we evaluate the effectiveness of satellite imagery-derived settlement maps in conjunction with various modeling techniques. We explore two estimation approaches based on settlement maps: a data-driven machine learning approach exemplified by a random forest model and a process-driven probabilistic approach exemplified by a hierarchical Bayesian model. Our findings underscore the efficacy of Bayesian modeling in addressing data scarcity and bias, providing robust estimates and quantifying model uncertainty. However, the random forest model performs better when data inputs are detailed and unbiased. We further emphasize the importance of considering settlement map characteristics in the modeling process, while recognizing the overall limitations of relying solely on satellite imagery for population counts. Through a rigorous evaluation of different stages of the population modeling pipeline—data input, model selection, and outcome assessment—this study provides key insights into the challenges and requirements of using satellite imagery-derived settlement maps for population estimation in data-scarce contexts.

1 | Introduction

On 15th November 2022, the world population reached "a milestone in human development" by crossing the threshold of 8 billion people (United Nations 2022). Behind this headline number, however, lies a disparity in the reliability of population figures across the world. Traditionally, population data have been collected through census exercises that require massive logistical and financial resources. In countries with political instability, conflicts, or natural disasters, logistical challenges can become insurmountable, resulting in outdated or incomplete data on population sizes and distribution for the most vulnerable regions. Consequently, 25% of countries in the world have national population totals that are more than 10 years old, with 12 countries where a last population count occurred before 2000 (United Nations Statistics Division 2023).

These population data inaccuracies distort the picture of a country's social and economic situation. For example, about half of the global sustainable development goal (SDG)

indicators involve a population denominator. Similarly, epidemiologists heavily rely on estimating the population at risk to study disease prevalence, a challenge referred to as the denominator problem (Garson 1976). Furthermore, inaccurate subnational population numbers hamper a fair allocation of resources within a country and across countries, which in return impedes efficient decision-making in many areas of public interest, such as urban planning, environmental hazard risk management and public health.

In recent years, the increasing availability of nontraditional data sources, from satellites to digital technologies, has diversified the data ecosystem for population research (Kashyap 2021). To what extent can this new data ecosystem support the estimation of current population sizes, especially in areas that are hard to access with conventional census-based data collection methods? A body of research dating back to the 1950s has highlighted the value of satellite imagery in producing population estimates (see Wu et al. (2005) for a comprehensive review). The utilization of geospatial data to estimate population sizes is itself not a novel concept. As early as 1780, extensive manual mapping was employed to derive population estimates of the French kingdom (Brian 2001). However, products derived from satellite imagery have revolutionized the availability of current, consistent, and detailed proxies for human settlements with comprehensive spatial coverage. Initially, satellite data were leveraged to produce intercensal population estimates (Iisaka and Hegedus 1982; Ogrosky 1975) or to enhance the spatial resolution of demographic data by spatially disaggregating census totals (Stevens et al. 2015; Leyk et al. 2019) which can both be characterized as census-dependent approaches. Nevertheless, a recent line of research has investigated whether satellite imagery can directly contribute to population estimation in contexts where no census data are available, an approach that has been termed as census-independent or bottom-up (Wardrop et al. 2018).

Drawing from a sample of small locations where the population has been entirely enumerated, the census-independent population estimation approach can further be categorized into two groups based on their use of satellite imagery. The first approach directly exploits very high-resolution raw satellite imagery and has been coined as a pixel-based approach (Mesev 2003). It relies nowadays on machine learning/deep learning algorithms such as support vector machines (Weber et al. 2018), representation learning (Neal et al. 2022) or the ResNet-18 model (Georganos et al. 2022). While promising, this study is constrained by the use of very high-resolution proprietary images (Vivid 2.0 from Maxar, WorldView Satellite and Pleiades) and by computer resources. As a result, this approach has been applied in a limited way so far to relatively small areas: two states in Nigeria in Weber et al. (2018), a sub-region of 850 km^2 in Neal et al. (2022) and three capital cities in Georganos et al. (2022).

The second approach leverages products derived from satellite imagery rather than raw satellite images, minimizing the need for demographers of direct knowledge in photogrammetry. Recent progress in computer vision has, for example, facilitated the creation of high-resolution settlement maps. These maps, encompassing all settlements within a given region, have been successfully integrated into countrywide population prediction models, as demonstrated in a study on Nigeria (Leasure et al. 2020). This hierarchical Bayesian population model has since been expanded to produce population estimates for Zambia (Dooley et al. 2021), for five provinces of the Democratic Republic of Congo (Boo et al. 2022), to complement the Burkina Faso census (Darin et al. 2022) and the Colombian census (Sanchez-Cespedes et al. 2024). There are several strengths in adopting a Bayesian approach for predicting population across an entire country using a small set of locations (below 1500 locations for Nigeria and the Democratic Republic of Congo). First, it is well-suited for data-scarce contexts where the lack of enumerated population data can be compensated by modeling the structure of the underlying population process and by providing robust estimates of model uncertainty. Second its flexibility allows for the integration of custom modeling subcomponents that address the challenges of different data contexts such as accounting for incomplete population enumeration (Dooley et al. 2021), accounting for population-weighted sampling designs (Leasure et al. 2021), breaking down population counts by age group and sex (Boo et al. 2022), and finally integrating observed building count (Sanchez-Cespedes et al. 2024). Efforts to enhance the technical accessibility of these methods have included developing openly available tutorials (Darin et al. 2021) and numerous workshops with national statistical offices, to address the need of multiple countries for up-to-date population estimates in the context of increasing barriers to ground data collection.

However, to date, despite increasing interest in census-independent approaches for low- and middle-income countries, the validity of the settlement map-based population modeling approach has never been assessed in a context where complete and reliable population data are available. The validation process conducted in the above-mentioned studies that use these methods has been based on out-of-sample cross-validation exercises involving the subdivision of a sample of locations into a training and a test set. This validation process reduces the size of an already small data set for training the model and is unable to assess population estimates at aggregated spatial scales. Furthermore the main satellite imagery derived product used across existing works is a proprietary building footprint layer produced by Ecopia AI and Maxar Technologies (Ecopia.AI & Maxar Technologies 2019). Since the first census-independent approaches however, several other institutions have released alternative openly-accessible settlement maps with global coverage and fine spatial resolution that have not yet been assessed for producing countrywide population estimates. Those settlement maps differ in terms of their objects (built-up area vs building), spatial details (10 m settled pixel vs building polygons), frequency of update (up to every 5 days) and producer type (private vs. public institutions).

In this study, we examine the potential of census-independent population modelling using satellite imagery-derived settlement maps in a uniquely controlled and reliable data environment provided by the 2018 Colombian Census. The Colombian Census data that we use in this study contain a census quality indicator, which identifies units with higher (over 95%) census coverage. The availability of this coverage

quality indicator, together with Colombia's wide range of environmental landscape, provides us a unique opportunity to assess the performance of population models derived from satellite-imagery at different stages of the population modelling pipeline. First, at the data input stage, we can study the impact of different settlement maps and thus help to design a strategy in a context of fast-evolving data products. For the ground truth population data, we can study the impact of sample sizes, which is a relevant consideration as satellite-derived population models are often deployed in contexts where ground data collection is often very resource intensive. Second, at the modelling stage, the controlled Colombian data environment allows us to compare two different modelling approaches, a probabilistic Bayesian model and a machine learning model to examine to what extent different modelling strategies can address limitations of the data input. Finally, at the outcome stage, we can assess accuracies in predicted population counts for different spatial scales and settlement environments. Together, this analysis advances a comprehensive, empirically informed understanding of the promises and pitfalls of estimating population sizes using satellite imagery-derived settlement maps in data-scarce contexts.

1.1 | Data

1.1.1 | Population Data

The 2018 Colombian census offers a high-resolution population data set, consisting of counts of the *de jure* (usual) population across 551,028 enumeration areas. We obtained access to the coverage assessment conducted by the census ground teams, which enabled us to select 319,837 enumeration areas considered complete, that is, where at least 95% of the population was covered. The goal of this process was to establish a controlled environment with highly reliable population data. The coverage assessment also helped identify areas that were difficult to survey. We classified the complete enumeration areas based on the mean coverage rate of the municipality and settlement type they belonged to. Specifically, we categorized enumeration areas as 'hard-to-reach' if the mean coverage rate for their municipality and settlement type fell below the 10th quantile of the overall mean coverage rate distribution, corresponding to 69% coverage. For comparison, the national coverage rate for the 2018 census was 91.5%.

In addition to these data-access considerations, Colombia presents an interesting context for studying the relationship between population and satellite-imagery derived products because of its diverse environmental landscape, as depicted in Figure 1. Seventy-seven percent of its population is concentrated in urban areas, particularly in the largest cities of Bogotá (the capital), Medellín, Cali, and Barranquilla. The central region of Colombia, *Central*, which encompasses the Andean Mountain range and the highland plateau, has a relatively high population density. This is primarily due to factors such as fertile agricultural land, economic opportunities in cities like Bogotá, and historical settlement patterns. The northern coast of Colombia, *Caribe*, is another region with a significant population size. Cities like Barranquilla have a dense population due to their coastal location, ports, tourism, and economic activities such as trade and industry. The Pacific coast of Colombia, *Pacifica*, located in the west, has a lower population density compared to the Caribbean coast. The region is characterized by dense rainforests, challenging terrain, and limited infrastructure, which has hindered large-scale settlements. The eastern plains, *Llanos/Orinoquia*, have a lower population density compared to other regions. This vast grassland and savannah area is primarily used for agriculture, cattle ranching, and oil production. The southeastern part of Colombia, which includes the Amazon rainforest, has a relatively low population density. The area is largely covered by dense tropical rainforests and is home to indigenous communities. Access to this region is often challenging, limiting large-scale settlements and has therefore a lower proportion of enumeration areas completely covered by the census as shown by Figure 1.

1.1.2 | Settlement Data

Early studies in population estimation from satellite imagery in data-scarce contexts relied on the manual digitization of building structures and rooftops, primarily using images from Digital Globe, for refugee settlements (Checchi et al. 2013), indigenous communities (Walker et al. 2014) and a city in Sierra Leone (Hillson et al. 2014, 2015, 2019). Advancements in computational power and machine learning systems for feature extraction, coupled with the increased availability of very high-resolution satellite images, have facilitated the automated production of countrywide settlement maps. Weber et al. (2018) leveraged this potential to estimate the population sizes in two Nigerian states from a sample of locations using satellite imagery to support vaccination campaign efforts. Due to the specialization of large institutes in settlement extraction models, subsequent years have seen a division between population modelers and settlement map producers. These include partnerships between Ecopia.AI and Maxar Technologies, the Joint Research Centre of the European Commission, the German Space Agency, Microsoft, and Google for the creation of new settlement layers.

The next generation of census-independent models has highlighted the valuable information contained in satellite-derived building footprints for population estimation from sampled locations (Boo et al. 2022; Darin et al. 2022; Dooley et al. 2021; WorldPop, & Institut National de la Statistique du Mali 2022; WorldPop, & National Population Commission of Nigeria 2021). All these models rely on a building footprint layer produced by Ecopia. AI & Maxar Technologies (2019). While a rasterized version of the layer was later made available covering the whole of sub-Saharan Africa (Dooley and Tatem 2020) the raw product is not openly available and no updates are planned for the foreseeable future. Other large-scale products are however openly available and regularly updated as described in Table 1.

Table 1 describes all the current openly available large-scale settlement maps and classifies them into three categories:

1. *Building footprints*: Building footprints are a polygon-based vector layer that automatically extracts the

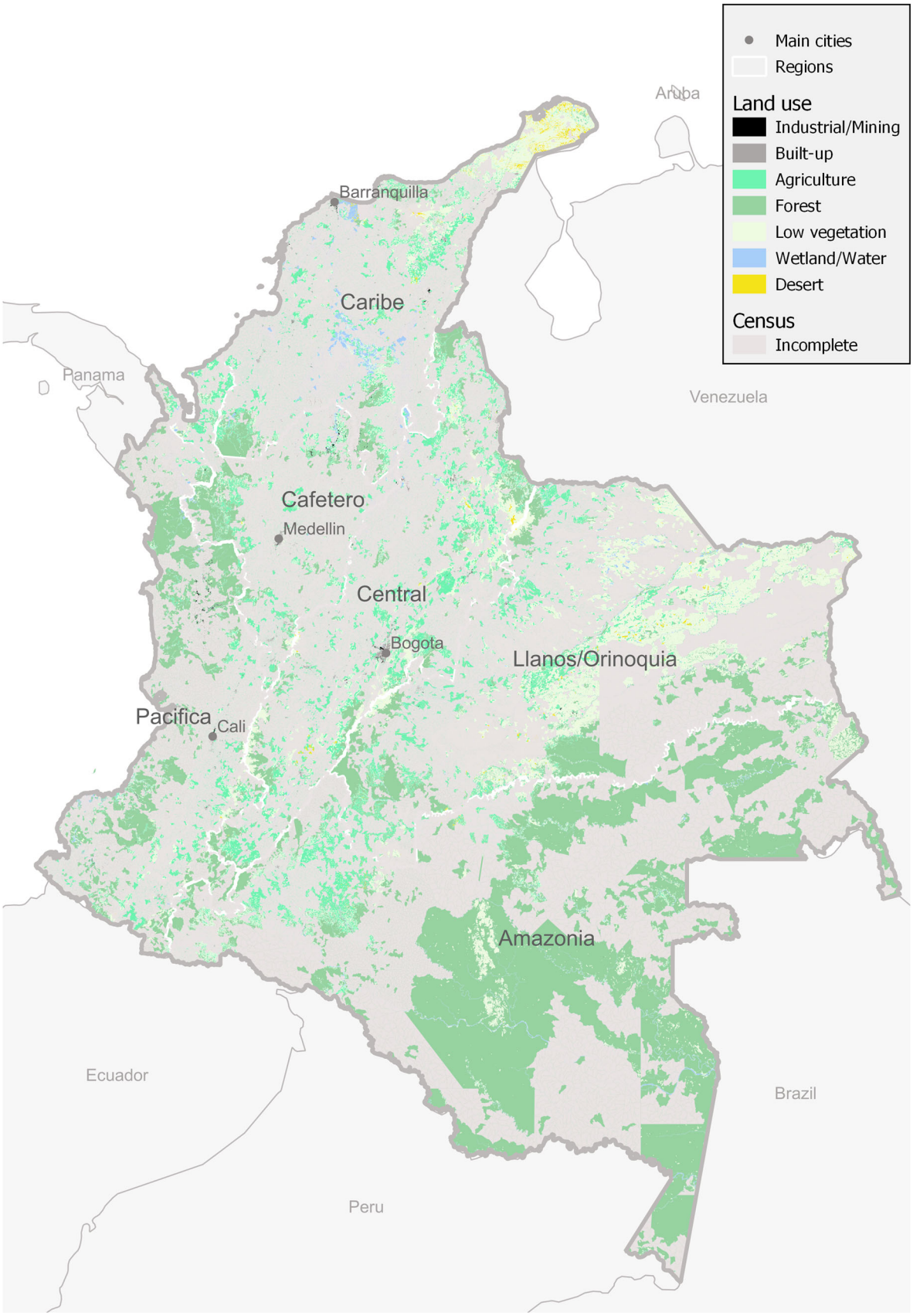

FIGURE 1 | Contextual information of the controlled data environment as provided by the Colombian 2018 census. This map represents the land cover information as provided by the Copernicus CORINE land cover reclassified in seven classes (industrial/mining, built-up, agriculture, forest, low vegetation, wetland/water, desert) for the enumeration areas of the census that have a coverage rate of 95% or more.

TABLE 1 | Settlement map attributes. The last column shows the percentage of missing enumeration areas per settlement type in Colombia with 4 the most remote and 1 the most urban.

	Attributes				Missing EAs per settlement type (%)			
	Format	**Resolution**	**Institution**	**Update**	**4**	**3**	**2**	**1**
Land cover								
Dynamic world	Raster	10 m	Private	Weekly	13	6	6	4
Built-up surface								
Global human settlement layer	Raster	10 m	Public	5-year	11	2	0	0
Word Settlement footprint	Raster	90 m	Public	ad hoc	18	7	3	1
Building footprint								
Microsoft	Polygon	50 cm	Private	ad hoc	15	8	4	1
Google	Polygon	50 cm	Private	ad hoc	24	11	7	5

delineation of every building. There are two large scale and openly accessible building footprint layers:

1. The first one is produced by Microsoft using Bing Maps imagery with a spatial resolution ranging from 30 cm in dense area to 50 cm in sparse areas and a period ranging from 2014 to 2022 (Bing Maps Team 2022).
2. The second is produced by Google as part of their Open Buildings v3 release that uses Google Maps imagery with a spatial resolution of 50 cm. No specific imagery date year is provided (Sirko et al. 2021).

2. *Built-up surface*: Built-up surface layers are pixel-based raster data sets that map the distribution of the built-up surface. We assessed two global and openly accessible built-up surface layers.

1. The first one is the World Settlement Footprint produced by the German Space Agency using imagery from Sentinel-1 and Sentinel-2 that has a spatial resolution of 90 m and exists for two time points: 2015 and 2019 (Marconcini et al. 2021). The second one is the Global Human Settlement Layer produced by the Joint Research Center using the same imagery and with a spatial resolution of 10 m. They produced an additional product that focuses only on the residential area that we consider as a separate layer in this study. Those layers exist also at 100 m resolution for every 5 years between 1975 and 2020 with two additional forecasts for 2025 and 2030 (Pesaresi and Politis 2022).

3. *Land cover map*: Land cover layers are a pixel-based raster data sets that classify the Earth's surface based on the type of cover (e.g., urban, forest, agriculture). Because the power to predict population sizes is potentially hampered by the coarse land cover classes identified by these products, we focus on products with high temporal resolution to see if this can improves their value for prediction. We assess the Dynamic World layer produced by Google in partnership with the National Geographic Society and the World Resources Institute, using imagery from Sentinel-1 and Sentinel-2 that have a 10 m spatial resolution with a temporal resolution of every 2 to 5 days (Brown et al. 2022). We processed it as the mean value per pixel over the 2018 year.

Figure 2 contrasts the six settlement maps to raw satellite imagery for several census blocks from the city of Neiva, Colombia. The main difference between them, as highlighted by the visual comparison, lies in their spatial resolution. To summarise settlement information for census enumeration areas, we used the count of building footprint centroids within each enumeration area and the proportion of each enumeration area covered by built-up surfaces.

1.1.3 | Other Contextual Data

The objective of a census-independent population model is to estimate for small, sampled areas the relationship between population counts and various covariates linked to population density. These covariates should have full coverage for the entire country such that the model can subsequently predict population counts for areas that have not been sampled.

We described large-scale variations of population density in Colombia with information about settlement types (4 levels), regions (6 levels) and departments (33 levels). The information regarding settlement types was obtained from the Global Human Settlement Layer Settlement Model, known as GHSL-SMOD (Schiavina et al. 2022). We reclassified this data set into four classes, representing a range from the least urbanized (4) to the most urbanized (1) settlement. We described small-scale variations in population density in Colombia with a set of 100 m gridded covariates averaged at the enumeration area level. We selected six covariates that are globally available:

- Human activity: mean night light data from the Visible Infrared Imaging Radiometer Suite, extracted for 2018 at 500 m (Elvidge et al. 2017).
- Accessibility: friction surface enumerating land-based travel speed with access to motorized transport from the Malaria Atlas Project, produced for 2019 at 1 km resolution (Weiss et al. 2018).

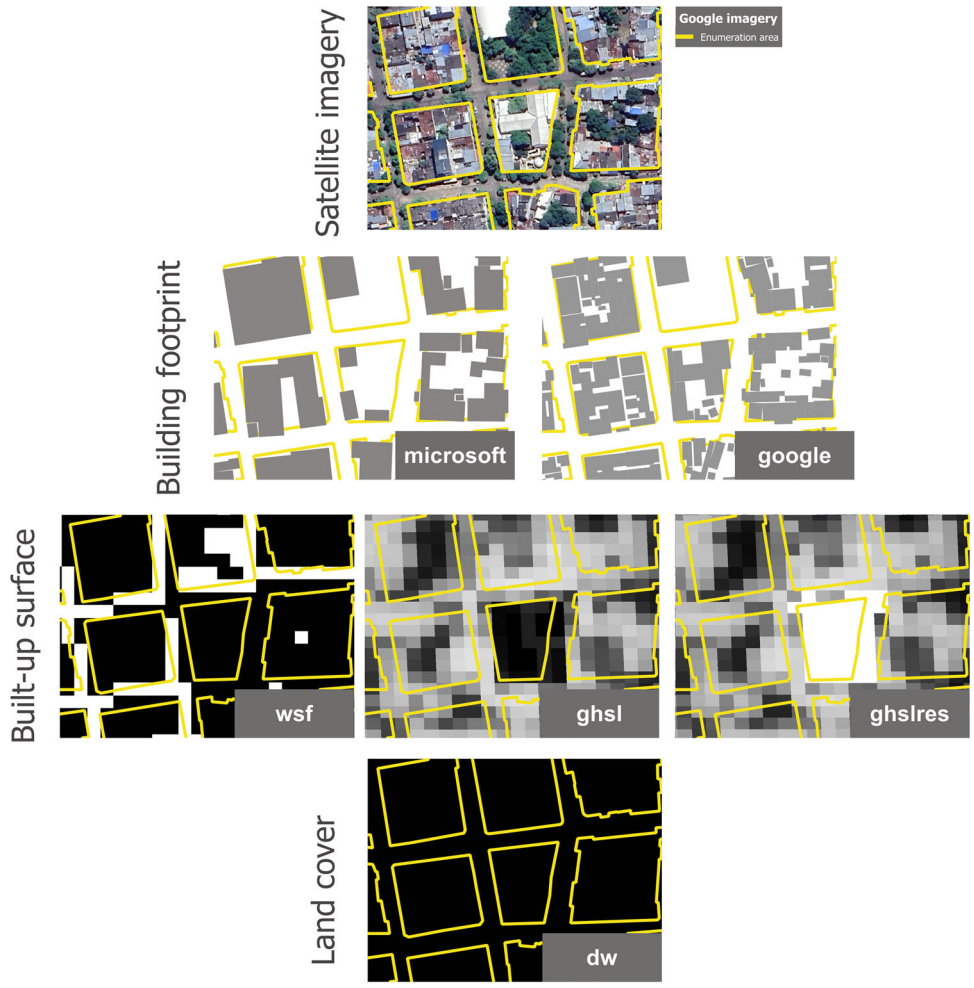

FIGURE 2 | Close-up on the six openly accessible global settlement maps within five census enumeration areas and ranked by spatial resolution.

- Urban layout: three covariates including the mean and standard deviation of building area in a 100 m window and the number of buildings in a 1 km window derived from the Microsoft building footprint layer (Bing Maps Team 2022).
- Building height: the average net height of built-up surfaces from the Global Human Settlement Layer for 2018 at 100 m resolution (Pesaresi and Politis 2022).

2 | Method

2.1 | Base Model: Hierarchical Bayesian Population Model

We used the hierarchical Bayesian population model from Leasure et al. (2020) as the base model for our analysis. The model's logic is demonstrated in Equation 2.1, where the observed population count (population_i) in location i was modelled with a Poisson distribution. The expected value of this count was determined by the product of two factors: settled area detected from satellite imagery (settled_area_i) and population density (λ_i^P) as a latent parameter. In Equation 2.2, λ_i^P was modelled using a log-normal distribution which incorporates unexplained variance σ in population densities, and the location parameter μ_i (i.e., log population density) was modelled as a linear regression with geospatial covariates. The random intercept was composed of a national parameter, α, and deviations from it represented by δ_t for each settlement type t, $\dot{\delta}_r$ for each region r, and $\ddot{\delta}_d$ for each department d. As predictors of small-scale variations in population density, the regression included a matrix of covariates X_i specific to each sample location i such as the average building height. Random effects by settlement type t and region r were estimated for each covariate.

$$\text{population}_i \sim \text{Poisson}\left(\lambda_i^P * \text{settled\_area}_i\right) \quad (2.1)$$

$$\lambda_i^P \sim \text{LogNormal}\left(\mu_i, \sigma\right) \quad (2.2)$$

$$\mu_i = \alpha + \delta_t + \dot{\delta}_r + \ddot{\delta}_d + \beta_{t,r} X_i \quad (2.3)$$

Minimally informative priors were defined for the intercept (Equation 2.4), the slope (Equation 2.5) and the variance (Equation 2.6).

$$\alpha \sim \text{Normal}(10, 10)$$

$$\delta_t \sim \text{Normal}(0, \tau)$$

$$\dot{\delta}_r \sim \text{Normal}(0, \tau)$$

$$\ddot{\delta}_d \sim \text{Normal}(0, \ddot{\tau})$$

$$\tau \sim \text{HalfNormal}(0, 10)$$

$$\dot{\tau} \sim \text{HalfNormal}(0, 10)$$

$$\ddot{\tau} \sim \text{HalfNormal}(0, 10) \quad (2.4)$$

$$\beta_{t,r} \sim \text{Normal}(0, 1) \quad (2.5)$$

$$\sigma \sim \text{HalfNormal}(0, 2) \quad (2.6)$$

2.2 | Base Model Calibrated With a Building Count Component

Directly integrating the settled_area as an input data obfuscates the fact that the settled area is imperfectly detected from satellite imagery and therefore is not a direct observation. To link the satellite imagery products with observed information, we integrated an additional set of ground observations, the building count reported by the census team. The building count data from the census was used for model fitting, but was then considered missing for the model prediction stage to reproduce a data context where the only complete data available comes from satellite imagery.

The flexibility of the Bayesian framework enabled a two-step modeling approach. In the first step (Equation 2.7), we extended the base model from Section 2.1 by deriving population counts not directly from the satellite-derived settled_area, but instead from a latent true building count, λ_i^B. As detailed in Equation 2.8, λ_i^B is estimated from the observed building_count in the census via a measurement error model to not propagate errors specific to the enumeration of building count into the population model. In the absence of data on the measurement error, we implemented a simple Poisson error structure after confirming that this improved out-of-sample population estimates and maintained good coverage of prediction intervals relative to using the observed building counts directly. The latent λ_i^B is then linked to the satellite-derived settled_areathrough a lognormal regression, whose mean includes a random intercept $a_{t,r}$ and slope $b_{t,r}$, both varying by settlement type and region. In effect, this calibrated population model shifts focus from modelling population per unit of settled area to modelling population per building.

$$\text{population}_i \sim \text{Poisson}\left(\lambda_i^P * \lambda_i^B\right)$$

$$\lambda_i^P \sim \text{LogNormal}(\mu_i, \sigma)$$

$$\mu_i = \alpha + \delta_t + \dot{\delta}_r + \ddot{\delta}_d + \beta_{t,r} X_i \quad (2.7)$$

$$\text{building\_count}_i \sim \text{Poisson}(\lambda_i^B)$$

$$\lambda_i^B \sim \text{LogNormal}(a_{t,r} + b_{t,r} \log(\text{settled\_area}_i), s) \quad (2.8)$$

The priors for the parameters of the sub-model from Equation 2.7 defining population count were the same as those used for the base model. For the building count sub-model from Equation 2.8, we defined minimally informative priors for the intercept, the slope and the variance (Equation 2.9).

$$a \sim \text{Normal}(0, 10)$$

$$b \sim \text{Normal}(1, 5)$$

$$s \sim \text{HalfNormal}(0, 2) \quad (2.9)$$

2.3 | Random Forest Model

We assessed how the performance of the process-driven probabilistic approach described above compared to that of a data-driven machine learning approach. We opted for the random forest framework (Breiman 2001) with the population density on the log scale as the response variable, in other words $\log(\text{population}_i/\text{settled\_area}_i)$. Random forests have been shown to perform similarly to other machine learning/deep learning approaches in out-of-sample prediction when the observed units are small (Metzger et al. 2022). To study each settlement map, we varied the denominator (i.e., settled_area_i) used to compute the population density while keeping the same selection of input covariates X used in the base model of Section 2.3.1 as well as including the variables describing the region, department and settlement type. In this framework, we did not integrate the building counts observed by the census team because they were considered as only available for a sample of locations and random forest cannot accommodate missing data.

2.4 | Model Evaluation

All models were trained within the R software (R Core Team 2022) using the Rstudio interface (Posit team 2023). To perform the Bayesian estimation of the model, we used the Stan software (Stan Development Team 2023) through the CmdStanR interface (Gabry and Češnovar 2023). We ran four Markov Chain Monte Carlo chains that included a burn-in period of 1500 iterations. We retained 1000 iterations following the burn-in period which were used for analysis. To assess model convergence, we used the default warning issuing from CmdStanR and checked that every parameter had an R-hat (the potential scale reduction factor, the weighted average of the between-chain and within-chain variances) below 1.1. To perform the random forest algorithm, we used the randomForest R package (Liaw and Wiener 2002) specifying 500 trees, a random sample of 2 covariates per tree (the number of covariates divided by three), and five observations for terminal nodes which correspond to the default settings in a regression framework.

To evaluate the performance of each model, we conducted out-of-sample cross-validation by comparing predicted population counts for all enumeration areas that were not sampled for model fitting. This allowed us to make claims about unseen areas and provides a more conservative estimate of model fit. In

cases where the settlement map failed to identify any settled areas within the enumeration area, we interpreted this as equivalent to predicting zero people. We then summarized the performance with three statistics based on the percentage error such that errors in sparsely populated enumeration areas have the same importance as errors in densely populated enumeration areas. First, we looked at *inaccuracy* as measured by the median absolute percentage error (MAPE) to assess how close the predictions are in general to the truth. We then looked at the *bias* as measured by the median percentage error (MPE), to assess if the predictions systematically overestimate or underestimate. Finally, we looked at the *imprecision* as measured by the standard deviation of the percentage error to assess if the size of the prediction error varied a lot across enumeration areas.

To stress-test the performance of the population model, we summarized the performance of the three models for each of the six settlement maps across different contexts (six regions and three settlement types) when randomly sampling 1000 observations. For each combination of model and settlement map, we created 10 data sets by randomly sampling from the full-coverage census to assess the sensitivity of the modelling results to sampling variations.

3 | Results

3.1 | The Bayesian Approach Performed Better With Less Informative Data

Figure 3 illustrates that building footprint layers consistently outperformed other settlement maps as the most reliable proxy for population sizes across all models and contexts, underscoring the importance of using settlement maps with the finest spatial resolution and specifically targeted at buildings. More precisely Google- and Microsoft-based models had a median inaccuracy of 41 and 43 people (i.e., 124% and 131% respectively) compared to 50 people (200%) for all the built-up surfaces. Conversely, the performance of Dynamic World, which covers all land cover types daily but lacks a specific focus on settlements, was consistently the weakest among all contexts and models (median inaccuracy of 62 people or 261%).

The comparison between modeling strategies highlights the advantage of the random forest approach for highly informative data sets but conversely the strengths of Bayesian modeling in contexts with limited information. In fact, for all settlement maps that were less detailed than building footprints, the calibrated Bayesian model outperformed the random forest model. By leveraging a sample of observed buildings, Bayesian modeling effectively converted the settlement map into building estimates, a quantity more closely related to population than settled area. Another notable strength lies in its ability to impose structural constraints, resulting in significantly less variation in prediction errors. Across settlement maps, the median imprecision of the calibrated Bayesian model was 91 people, while the random forest approach exhibited a higher median imprecision of 167 people. This difference suggests that the calibrated Bayesian model is less sensitive to extreme values in the sample.

3.2 | The Bayesian Approach Performed Better With Biased Data

We looked at municipality prediction to see how well census-independent population models perform for larger spatial units. As expected, Figure 4 shows that aggregated population estimates were less prone to modelling error. The mean inaccuracy across all contexts, models and settlement maps dropped from 174% for enumeration area-level predictions to 40% for municipality-level predictions.

Assessing model performance for municipalities (i.e., aggregations of enumeration areas) revealed a prediction bias undetected at enumeration area level (see the left panel of Figure 4). Models using building footprints, particularly Microsoft, underestimated population counts for the municipalities situated in the region of Pacifica and Amazonia where population is less dense (Section 1.1) and enumeration areas without detected building footprints are more present (Table 1). The calibrated Bayesian model that utilized observed building count data better accounted for the under-detection of buildings with a median bias of −17% compared to −36% for the random forest model in the Amazonia region. The bias directly impacted on model performance at the municipality level. The best model was the calibrated Bayesian model for each of the settlement maps ranging from an average inaccuracy of 31% with the Global Human Settlement layer to 37% with Dynamic World. The distribution of the inaccuracy metrics across ten samples as displayed on the left panel of Figure 4 shows unexpectedly that the random-forest model did not lead to more variations between samples than the calibrated Bayesian model (a standard deviation of 10 percent point). A last insight from Figure 4 lies in the poor performance of the base Bayesian model that showed a median positive bias of 83% and inaccuracy of 95%.

3.3 | Sparsely Populated Areas Were Harder to Predict

The municipality level assessment because of its larger size enabled us to account for the prediction bias seen in Figure 3 but obfuscated variation of prediction accuracy between settlement types. We thus created a custom municipality breakdown by settlement type by aggregating up the enumeration area predictions for which the census observations are complete per municipality and settlement type—a geographical scale that we named 'medium scale'. The results presented in Figure 5 illustrate that across all models and settlement maps, prediction inaccuracy was greatest for very sparsely populated areas as shown by a mean inaccuracy of 50% compared to 29% for dense areas. This difficulty arises from the challenge in detecting remote settlements using satellite imagery products (Sanchez-Cespedes et al. 2024). The dense and very dense settlement types of all regions are the easiest to capture, except Central that contains the capital Bogotá. This may be due to higher heterogeneity in building occupancy across settlement types in the Central region with 72% residential buildings in very sparse areas and 87% in very dense areas according to the census. Generally, prediction inaccuracy is negatively correlated with the proportion of residential buildings (−0.21) and positively

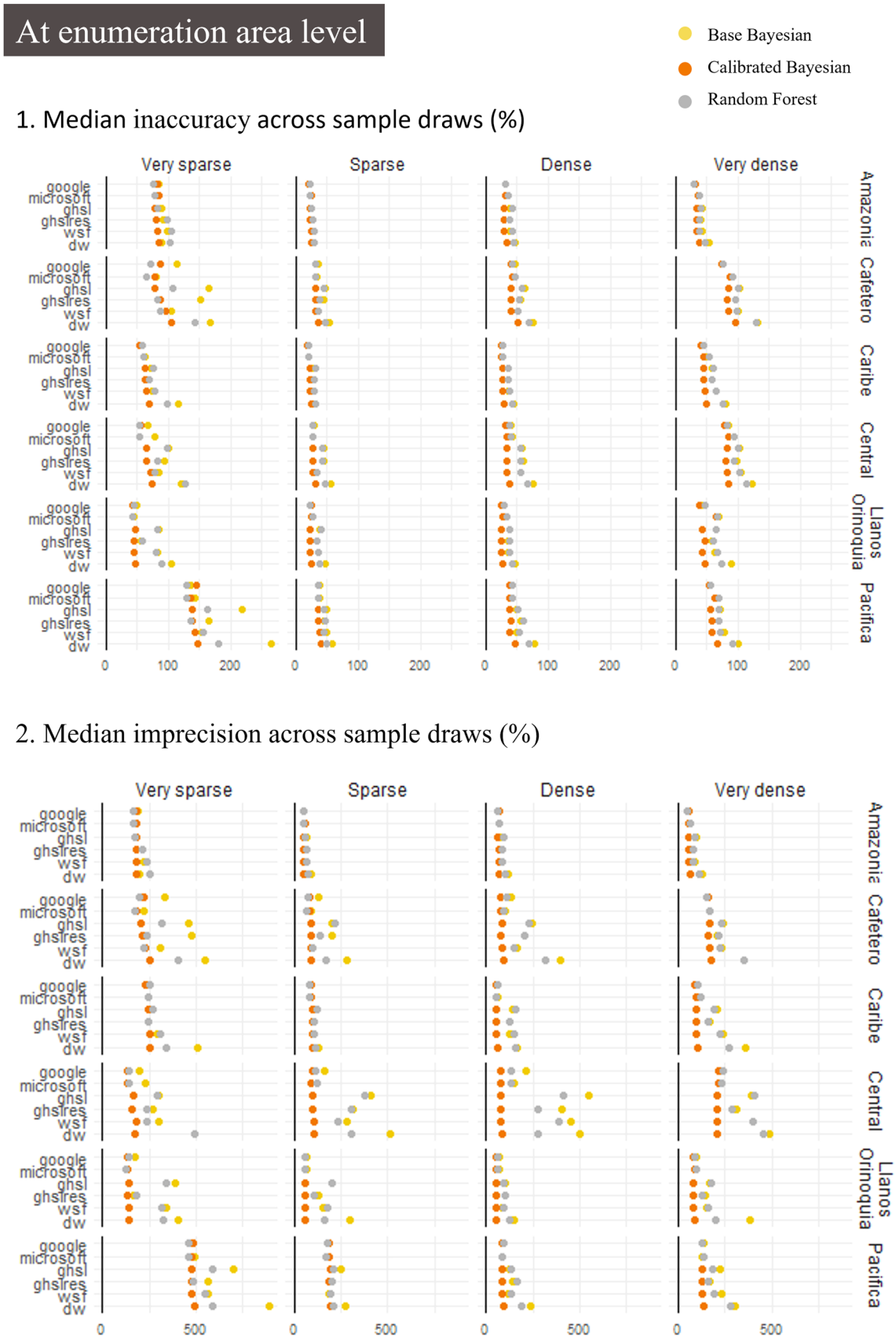

FIGURE 3 | Comparison of performance across settlement maps and models at small scale for the four settlement types and six regions.

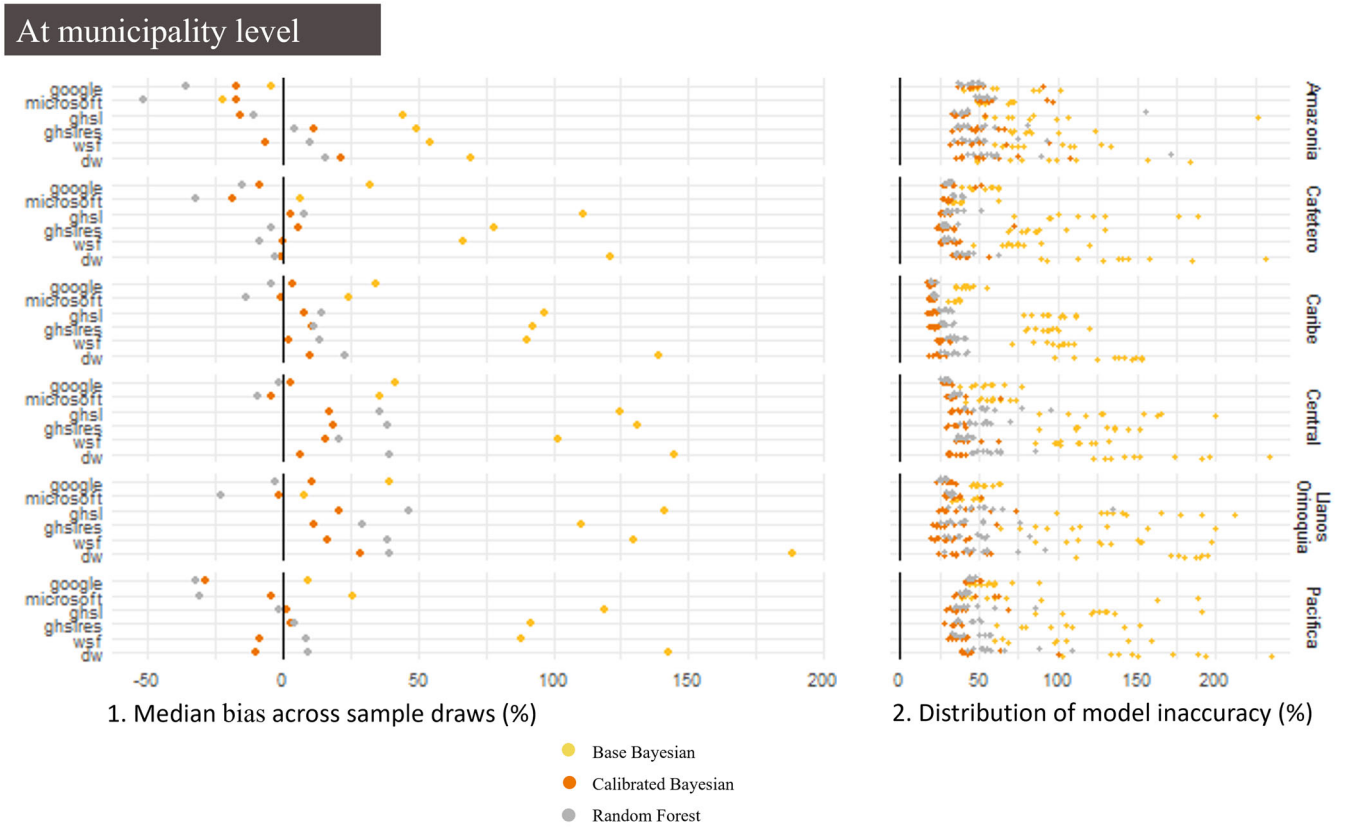

FIGURE 4 | Comparison of performance across settlement maps and models at the municipality level. 1. Median bias across sample draws (%). 2. Distribution of model inaccuracy (%).

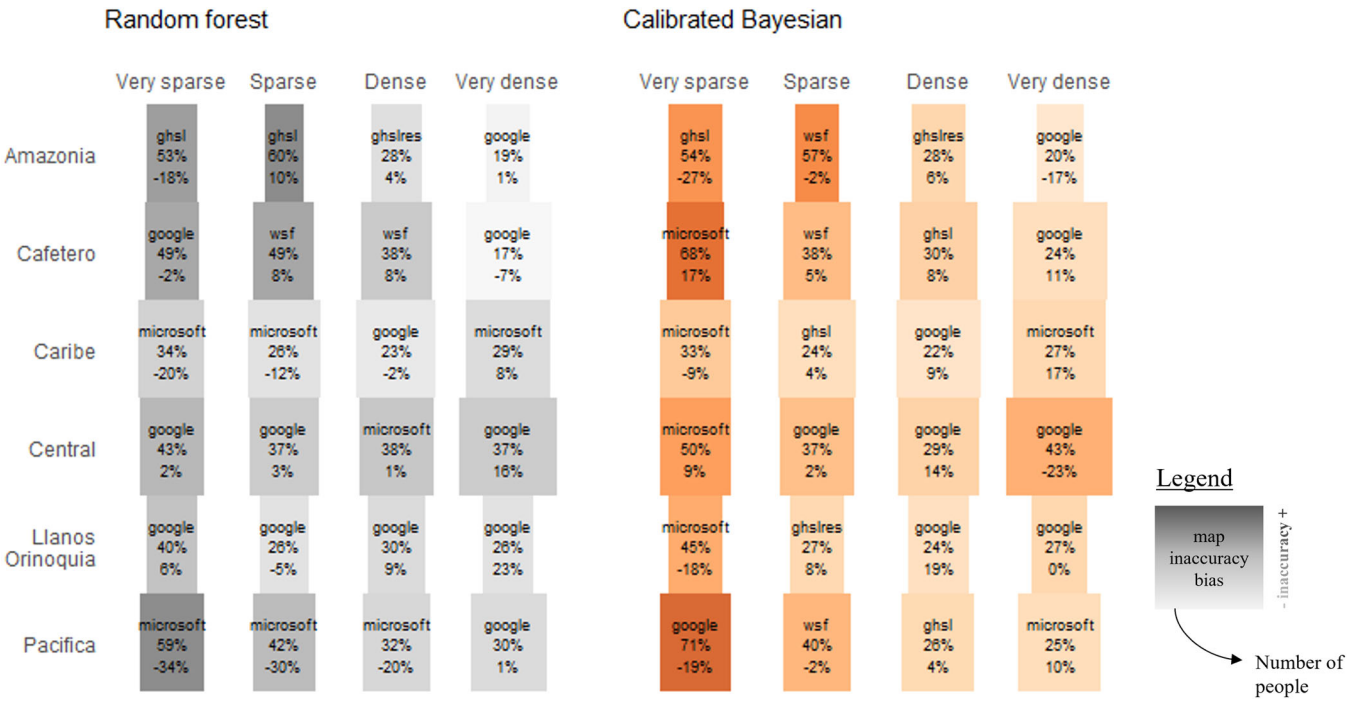

FIGURE 5 | Performance of the best random forest and calibrated Bayesian model per region and settlement type at medium spatial scale. The bigger the box the more the people, the darker the box the higher the inaccuracy.

correlated with the standard deviation of the proportion of residential buildings (0.29).

Revisiting the assessment of settlement maps, we see that the building footprint layers performed the best, with 34 of the 48 combinations of models and settings being more accurately predicted using the Google or Microsoft settlement map. This was especially the case when integrated within the random forest model where population counts were better predicted in 19 of the 24 settings with a building footprint layer (10 with Google and 9 with Microsoft). This result does not hold however in the remote Amazonia and Pacifica regions where, across models, the two layers had lower prediction accuracy than the built-up surface maps. Furthermore, the calibrated Bayesian model had better performances than the random forest one in 14 of the 24 settings, based on a mixture of settlement maps.

3.4 | More Samples Decreased the Sensitivity to Sampling Errors

Settlement map-based population models are adopted in settings where ground data on population counts is hard to collect. It is therefore key to understand how increasing the sample size of ground truth data impacts on model performance.

Figure 6 shows how increasing the sample size from 500 to 10,000 enumeration areas had a small positive effect on the accuracy of the population model (from 224% to 202% inaccuracy at enumeration area level and from 44% to 33% at municipality level). This average accuracy increase was mainly led by the performance gain from the coarser settlement maps. The model that was the most impacted by greater sample size was the random forest model that saw enumeration area level prediction inaccuracy decrease by 31% points on average.

The greatest impact however of a larger sample size as depicted in the left panel of Figure 6, was the reduced variability of model performance across different sample draws, as indicated by the drop in the standard deviation of the inaccuracy metric computed for each of the 10 sample draws. Indeed, the standard deviation reduced by 96% points on average across all the combination of models and settlement maps at enumeration area level (32 percentage points at municipality level). Increasing the sample size thus reduces the sensitivity to sampling error and increases the consistency with which models reached the best prediction possible.

3.5 | Population Estimates Were Less Accurate in Densely Forested Areas

Figure 7 classifies the medium spatial scale units according to their main land use (Figure 1). It shows a trend that is common across the two population models: better performance was achieved in a built-up context than in rural areas. Because there are more spatial units in the built-up context (51%), change in accuracy will have a larger impact on countrywide accuracy assessment. The areas that were the most difficult to predict were by contrast the densely-forested areas of Amazonia and Pacifica.

The last analysis assessed if population estimation can fill gaps where the census enumeration had the lowest coverage. On

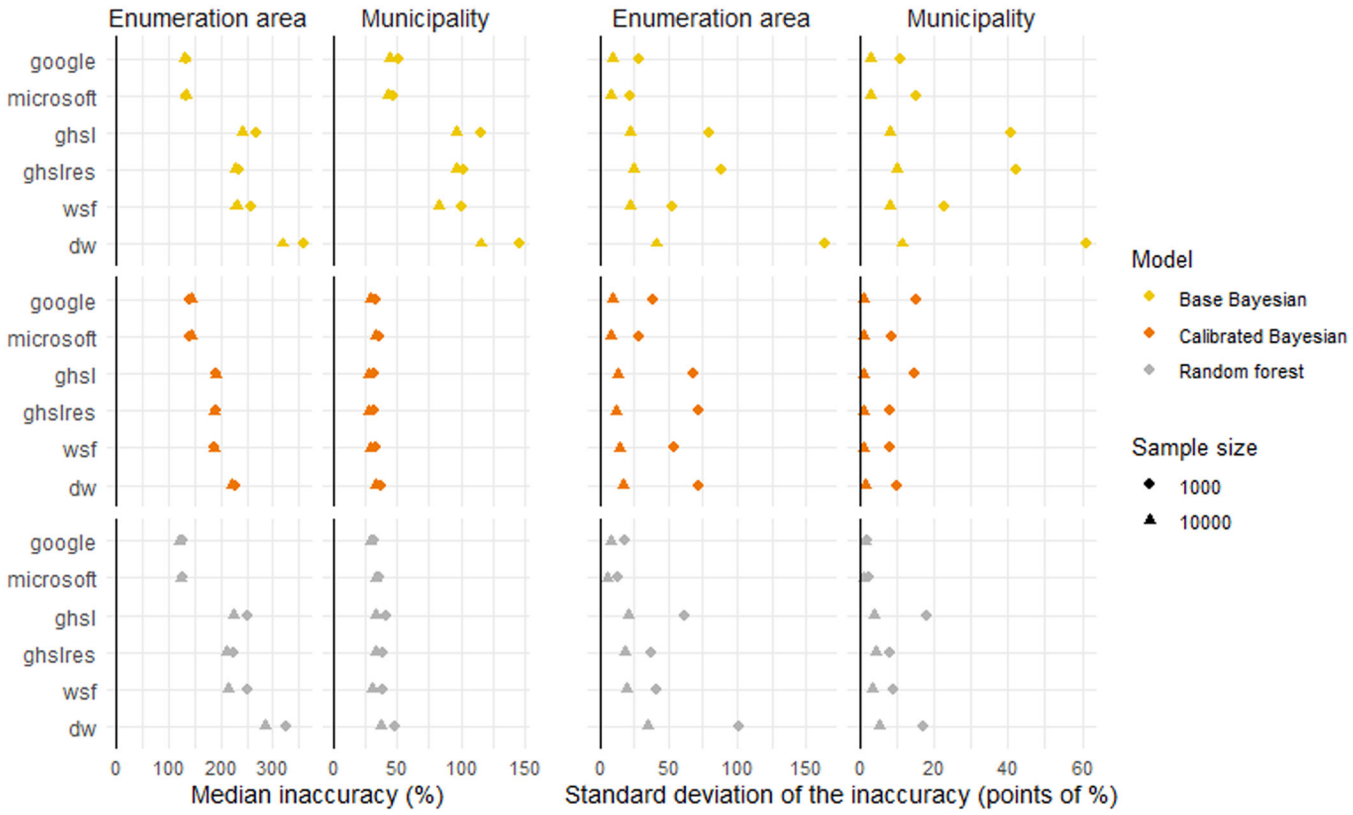

FIGURE 6 | Comparison of models performance across settlement maps and models for different sample sizes at small and large spatial scale.

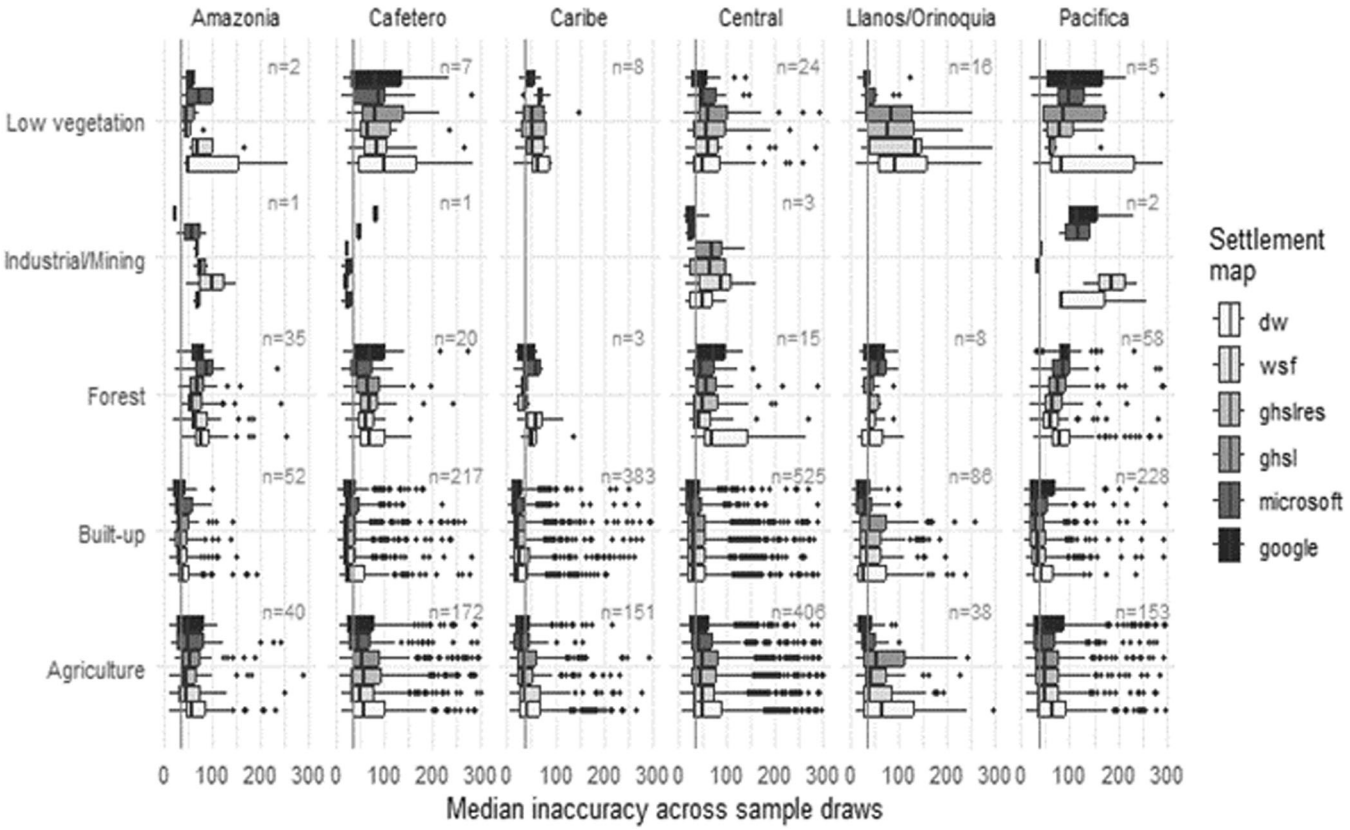

FIGURE 7 | Distribution of the median inaccuracy across the 10 sample draws at medium spatial scale per land use (Axis truncated at 300% inaccuracy). The vertical line shows the median inaccuracy across models (35.4%). The number of spatial units in each land use class is displayed on the left-hand side of the plot.

average, population prediction for accessible enumeration areas had an average inaccuracy of 56% when for inaccessible enumeration areas it reaches 91%. Figure 8 illustrates that the discrepancy is larger for coarse settlement maps (between 38% and 55% point difference) than for building footprint settlement maps (17 for Microsoft, 29 for Google). But the major difference is between regions: Figure 8 shows that in Amazonia the population hard to enumerate with traditional door-to-door collection methods is as challenging as for model-based estimation methods (186 percentage point difference).

4 | Discussion

Using Colombian census data and a detailed coverage assessment, we created a controlled environment of high-quality population data across a varied range of landscapes to assess the performance of two types of satellite imagery derived population models—process-driven probabilistic or data-driven machine learning—for different spatial scales, settings and settlement maps. Although our analysis is specific to Colombia, our findings can help inform population estimation strategies in data-scarce regions globally.

4.1 | The Importance of Building Counts for Estimating Population

The strong correlation between building counts and population sizes has fueled a recent surge of interest in harnessing satellite-derived settlement maps for population estimation. The initial endeavor to derive population figures from satellite images in data-scarce settings emerged in the context of temporary refugee settlements, relying on manual counts of visually identified building-like features (Checchi et al. 2013). Subsequent progress in computer vision and feature-extraction algorithms, coupled with increased spatial resolution of satellite imagery, led to the creation of large-scale automated settlement maps. While the past two decades have seen the global release of openly available built-up area maps as pixel-based representations of building coverage or settlement footprints (Marconcini et al. 2021; Pesaresi et al. 2013; Pesaresi and Politis 2022), the last 5 years have witnessed a proliferation of deep learning methods and computational power, enabling the extraction of built-up maps as polygon-based representations of building footprints (Bing Maps Team 2022; Ecopia.AI & Maxar Technologies 2019; Sirko et al. 2021). This wealth of information within building footprint maps has spurred the development of several population estimation models (Boo et al. 2022; Darin et al. 2022; Leasure et al. 2020; Metzger et al. 2022), and even a sub-Saharan Africa-wide interactive application to produce custom rapid response population estimates based on multiplying building counts by user-defined population ratios (Leasure et al. 2020).

Currently, openly available, large-scale building footprint products are offered by Google (Sirko et al. 2021), Microsoft (Bing Maps Team 2022) and OpenStreetMap (OpenStreetMap contributors 2023). However, the OpenStreetMap building layer has already been identified as problematic for country-wide mapping in data-scarce environments due to unequal coverage of manual mapping

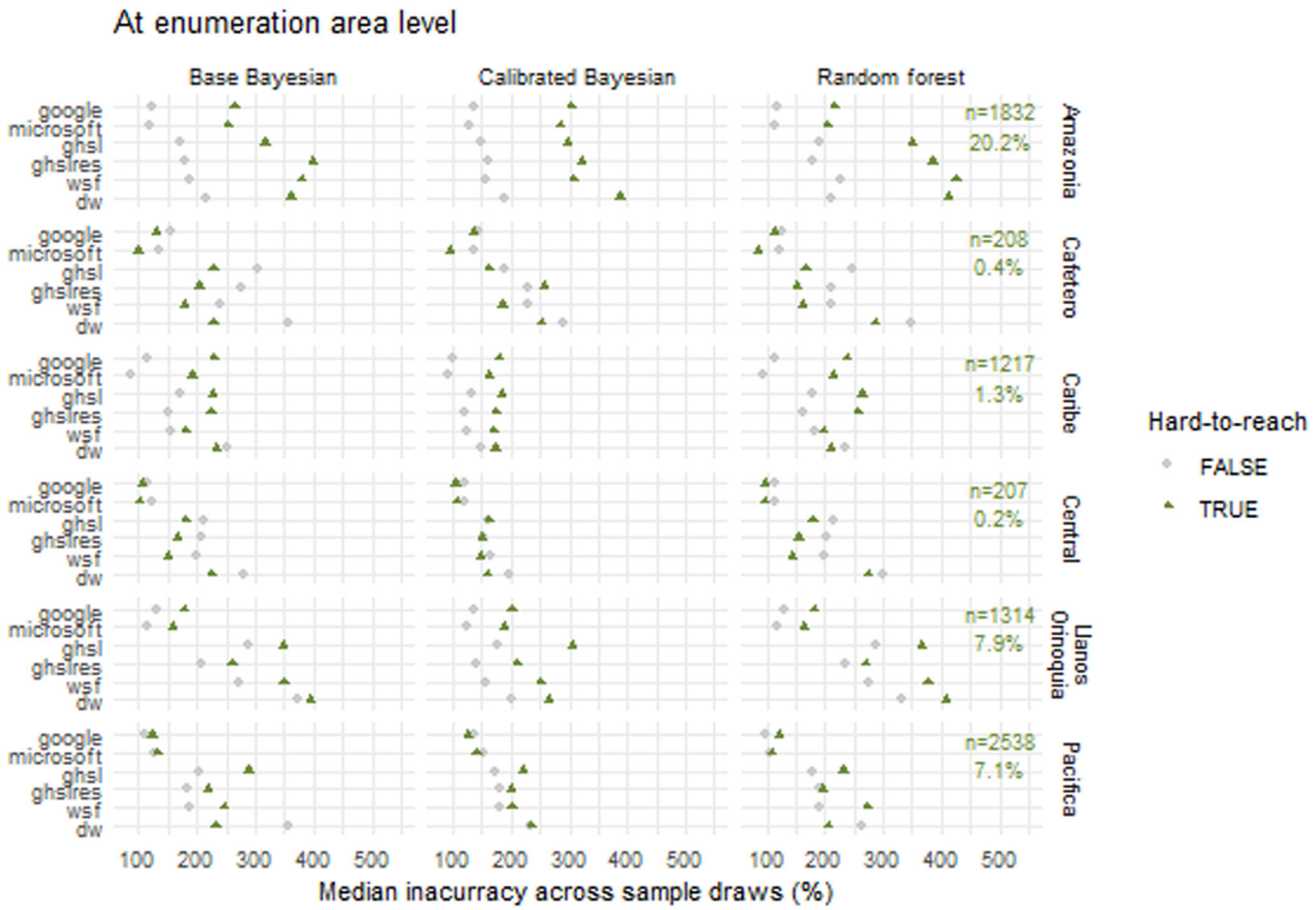

FIGURE 8 | Comparison of the median inaccuracy at enumeration area level across settlement maps, regions, and settlement types in the hard-to-reach areas. In green is indicated the number and the percentage of enumeration areas that are defined as hard to reach per region and settlement type.

efforts (Chamberlain et al. 2024). Our study demonstrated that the other two building footprint layers, Google and Microsoft, tend to outperform pixel-based settlement maps across both model types. This aligns with Hillson's (2019) findings indicating that building counts are superior to rooftop area for population estimation as illustrated for the city of Bo in Sierra Leone. Continued improvements in open-source building detection from remote sensing therefore provides distinct advantages for population modeling.

However, we observed a systematic large negative bias in population estimation for two remote regions (Amazonia and Pacifica) using both building footprint layers. This highlights disparities in prediction accuracy in specific settings where settlements are harder to capture potentially due to less conventional housing structures or thicker forest cover, or in the case of Microsoft omitting to process areas considered as inhabited, and it challenges the notion of building footprints as the default settlement map for population estimation. Additionally, both data layers are produced by private companies with irregular and unpredictable open release timelines, while other producers, such as Ecopia AI, offer custom building footprint products commercially.

Nonetheless, our study demonstrates that in the absence of a reliable building footprint layer, leveraging building count observations collected on the ground enables the conversion of coarse settlement information into an estimate of building counts which results in higher accuracy of population estimates. This underscores the importance of integrating building listings into routine ground data collection during household surveys and making these building listings available to inform both population and building modeling. It is noteworthy that observed building counts do not need to originate from the same data source as observed population counts, highlighting the flexibility of the modeling framework. This finding aligns with the lessons learned from integrating social cartography in population estimation for settlements that are harder to detect in satellite imagery, such as indigenous communities (Sanchez-Cespedes et al. 2024). Finally, basing population models on the presence of settlements overlooks subpopulations that do not live-in buildings, such as the homeless or nomadic groups, which will lead to underestimation of the number people especially if those subgroups are clustered in space.

4.2 | The Importance of Process-Driven Probabilistic Modelling in Data-Scarce Contexts

We have shown in this study that when the input data is complete, less biased, and detailed enough even off-the-shelf random forest models performed well. However, Bayesian probabilistic models performed better in contexts where the settlement layer is less informative because they can correct the

bias. This is by extracting more information from coarser settlement maps by leveraging sparse information such as sampled building counts to increase the accuracy of the prediction model. All in all, when articulating different settlement maps for the various regional and settlement type settings as shown by Figure 5, the calibrated Bayesian model had slightly lower inaccuracy (31.4%) than the random forest approach (35.3%). The calibrated Bayesian model also demonstrated good properties in terms of precision of the population estimates with lower imprecision and better stability across sample draws than the random forest approach at the enumeration area and at the municipality levels.

This illustrates how specific challenges of scarce data settings may not be overcome by purely data-driven machine learning techniques as represented by the random forest approach, which is sensitive to sampling errors in small samples. Two features of the training sample are indeed driving machine learning performance: how large it is and how exchangeable with the prediction space it is. We conclude that in data sparse settings we cannot assume that the sampled training data is representative and unbiased and thus we need to develop probabilistic models that posit the structure of the underlying process of population spatial distribution. Another advantage of adopting a probabilistic model is its capacity to quantify the uncertainty within the modelled population estimates. To provide a relevant comparison with the machine learning model, however, we have only presented in this study point estimates from the Bayesian probabilistic models.

Thanks to the controlled environment provided by the Colombian census, we were able to systematically assess the bias and inaccuracy of each model by comparing estimates with the true enumerated population at different spatial scales. In practice satellite imagery-based population models are used when little is known about the true size of the population. In these settings we recommend the following steps to validate the outputs: (1) triangulating with external complete population data sources at coarser scale such as the World Population Prospect (United Nations Department of Economic and Social Affairs, Population Division 2024) to ensure large-scale validity, (2) triangulating with external partial population data sources at finer scale such as provided by the health and demographic surveillance systems (HDSS) (Herbst et al. 2021), and (3) the gold standard of out-of-sample cross-validation to assess prediction accuracy in contexts relevant to each use case (e.g. predictions into new regions, filling gaps between sample locations, prediction for new times).

4.3 | A Demographic Data Revolution in Population Estimation?

The best performing model-settlement map combination, which paired the calibrated Bayesian model with the Google building footprint layer, achieved a median inaccuracy of 32% and a median bias of 1.5% at municipality level. This is a promising result in areas where no other data are available on population count. Our analysis has however highlighted the heterogeneity of the inaccuracy distribution. When decomposing the performance across subnational regions, we see that the median inaccuracy hides large disparity in the distribution of the prediction inaccuracy that ranges from 17% in Caribe in dense and sparse settings to 100% in the very sparsely settled areas of Pacifica which may be due to reduced accuracy of feature extraction algorithms for identifying buildings or built-up areas in densely forested areas (Muro et al. 2020; Sanchez-Cespedes et al. 2024). Similarly, the inaccuracy increased at finer spatial scales reaching a level of 148% for the same model combination. This study therefore illustrates the trade-off between spatial precision and reliability of the predictions, between densely and sparsely built environment and between high and low levels of mixed-use buildings.

Increasing the sample size of observations improves the precision of estimates, but has a smaller effect on their accuracy, highlighting the strong limitations of current models in relating satellite imagery-derived products to population counts using sparse survey data. To address these shortcomings, we suggest two pathways for improving accuracy across different settings and spatial scales. First, incorporating building occupancy types (e.g., residential, commercial, mixed, industrial) into the models could enhance the relationship between population and building counts, particularly in areas with many nonresidential buildings or high heterogeneity. Second, combining multiple settlement maps into a single model could improve performance by compensating for the shortcomings of individual data sources, addressing a key current challenge in demographic data revolution.

5 | Conclusion

The detailed 2018 Colombian census, along with the census coverage assessment, provided a unique environment to evaluate settlement map-based census-independent population models. At the data input stage, we found that building footprint layers are the most promising form of satellite imagery derived product for population models. At the modeling stage, our analysis showed that a data-driven machine learning approach such as random forests are the best to extract information from detailed and unbiased settlement maps. Our findings emphasize however the importance of a Bayesian approach when dealing with coarse or biased settlement maps, particularly in sparsely populated regions like Amazonia and Pacifica where collecting observed building counts helped address these challenges. At the outcome stage, we discovered that predictions are more reliable at an aggregated scale and that population models may perform poorly in remote areas. This highlights the need for methodological advancements to integrate diverse data sources for improved estimation accuracy.

Acknowledgements

The authors would like to express their gratitude to the Colombian National Administrative Department of Statistics (DANE) for providing secured access to confidential data that was crucial for this study. We particularly appreciated the efforts of Andres Felipe Copete Martinez, Dr Lina Maria Sanchez Cespedes, Dr Javier Sebastian Ruiz Santacruz, Associate Prof Leonardo Trujillo Oyola, and Directors, Associate Prof Piedad Urdinola Contreras and Prof Juan Daniel Oviedo Arango in

providing substantial support throughout our research project. Edith Darin gratefully acknowledges the resources provided by the International Max Planck Research School for Population, Health and Data Science (IMPRS-PHDS). This project was approved by the Departmental Research Ethics Committee of the Department of Sociology at the University of Oxford (SOC_R2_001_C1A_22_32).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available on OSF at https://osf.io/p594q/. Please note that no references to spatial location can be found for confidentiality purposes. Additionally, a sample of the code used for analysis is available the same location. Any further requests for data or code should be directed to the corresponding author.

References

Bing Maps Team. 2022. "Microsoft Has Released New and Updated Building Footprints." *Microsoft Bing Blogs*, January 14. https://blogs.bing.com/maps/2022-01/New-and-updated-Building-Footprints.

Boo, G., E. Darin, D. R. Leasure, et al. 2022. "High-Resolution Population Estimation Using Household Survey Data and Building Footprints." *Nature Communications* 13, no. 1: 1330. Article 1. https://doi.org/10.1038/s41467-022-29094-x.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45, no. 1: 5–32. https://doi.org/10.1023/A:1010933404324.

Brian, É. 2001. "Nouvel Essai Pour Connaître La Population Du Royaume.Histoire Des Sciences, Calcul Des Probabilités Et Population De La France Vers 1780." *Annales de démographie historique* 102, no. 2: 173–222. https://doi.org/10.3917/adh.102.0173.

Brown, C. F., S. P. Brumby, B. Guzder-Williams, et al. 2022. "Dynamic World, Near Real-Time Global 10 M Land Use Land Cover Mapping." *Scientific Data* 9, no. 1: 251. https://doi.org/10.1038/s41597-022-01307-4.

Chamberlain, H. R., E. Darin, W. A. Adewole, W. C. Jochem, A. N. Lazar, and A. J. Tatem. 2024. "Building Footprint Data for Countries in Africa: To What Extent Are Existing Data Products Comparable?" *Computers, Environment and Urban Systems* 110: 102104. https://doi.org/10.21203/rs.3.rs-3334423/v1.

Checchi, F., B. T. Stewart, J. J. Palmer, and C. Grundy. 2013. "Validity and Feasibility of a Satellite Imagery-Based Method for Rapid Estimation of Displaced Populations." *International Journal of Health Geographics* 12, no. 1: 4. https://doi.org/10.1186/1476-072X-12-4.

Darin, E., M. Kuépié, H. Bassinga, G. Boo, A. J. Tatem, and P. Reeve. 2022. "The Population Seen From Space: When Satellite Images Come to the Rescue of the Census." *Population* 77, no. 3: 437–464. https://doi.org/10.3917/popu.2203.0467.

Darin, E., D. R. Leasure, and A. J. Tatem. 2021. *Statistical Population Modelling For Census Support*. WorldPop, University of Southampton. https://doi.org/10.5281/zenodo.5572490.

Dooley, C. A., H. Chamberlain, D. R. Leasure, G. Membele, A. Lazar, and A. J. Tatem. 2021. *Description of Methods for the Zambia Modelled Population Estimates From Multiple Routinely Collected and Geolocated Survey Data, version 1.0*.

Dooley, C. A., and A. J. Tatem. 2020. *Gridded Maps of Building Patterns Throughout Sub-Saharan Africa, Version 1.0*. WorldPop Research Group, University of Southampton. https://doi.org/10.5258/SOTON/WP00666.

Ecopia.AI & Maxar Technologies. 2019. Digitize Africa Data. http://digitizeafrica.ai.

Elvidge, C. D., K. Baugh, M. Zhizhin, F. C. Hsu, and T. Ghosh. 2017. "VIIRS Night-Time Lights." *International Journal of Remote Sensing* 38, no. 21: 5860–5879. https://doi.org/10.5555/3141742.3141745.

Gabry, J., and R. Češnovar. 2023. *cmdstanr: R Interface to'CmdStan'* [Computer Software]. https://mc-stan.org/cmdstanr.

Garson, J. Z. 1976. "The Problem of the Population at Risk." *Canadian Family Physician Medecin de famille canadien* 22: 71–74.

Georganos, S., S. Hafner, M. Kuffer, C. Linard, and Y. Ban. 2022. "A Census From Heaven: Unraveling the Potential of Deep Learning and Earth Observation for Intra-Urban Population Mapping in Data Scarce Environments." *International Journal of Applied Earth Observation and Geoinformation* 114: 103013. https://doi.org/10.1016/j.jag.2022.103013.

Herbst, K., S. Juvekar, M. Jasseh, et al. 2021. "Health and Demographic Surveillance Systems in Low- and Middle-Income Countries: History, State of the Art and Future Prospects." Supplement, *Global Health Action* 14, no. S1: 1974676. https://doi.org/10.1080/16549716.2021.1974676.

Hillson, R., J. D. Alejandre, K. H. Jacobsen, et al. 2014. "Methods for Determining the Uncertainty of Population Estimates Derived From Satellite Imagery and Limited Survey Data: A Case Study of Bo City, Sierra Leone." *PLoS One* 9, no. 11: e112241. https://doi.org/10.1371/journal.pone.0112241.

Hillson, R., J. D. Alejandre, K. H. Jacobsen, et al. 2015. "Stratified Sampling of Neighborhood Sections for Population Estimation: A Case Study of Bo City, Sierra Leone." *PLoS One* 10, no. 7: e0132850. https://doi.org/10.1371/journal.pone.0132850.

Hillson, R., A. Coates, J. D. Alejandre, et al. 2019. "Estimating the Size of Urban Populations Using Landsat Images: A Case Study of Bo, Sierra Leone, West Africa." *International Journal of Health Geographics* 18, no. 1: 16. https://doi.org/10.1186/s12942-019-0180-1.

Iisaka, J., and E. Hegedus. 1982. "Population Estimation From Landsat Imagery." *Remote Sensing of Environment* 12, no. 4: 259–272. https://doi.org/10.1016/0034-4257(82)90039-6.

Kashyap, R. 2021. "Has Demography Witnessed a Data Revolution? Promises and Pitfalls of a Changing Data Ecosystem." *Population Studies* 75, no. sup1: 47–75. https://doi.org/10.1080/00324728.2021.1969031.

Leasure, D. R., C. A. Dooley, B. Maksym, and A. J. Tatem. 2020. *peanutButter: An R Package to Produce Rapid-Response Gridded Population Estimates From Building Footprints, Version 0.1.0*. WorldPop Research Group, University of Southampton. https://doi.org/10.5258/SOTON/WP00667.

Leasure, D. R., C. A. Dooley, and A. J. Tatem. 2021. *A Simulation Study Exploring Weighted Bayesian Models to Recover Unbiased Population Estimates From Weighted Survey Data*. WorldPop, University of Southampton. https://data.worldpop.org/repo/docs/leasure2021simulation/leasure2021simulation.pdf.

Leasure, D. R., W. C. Jochem, E. M. Weber, V. Seaman, and A. J. Tatem. 2020. "National Population Mapping From Sparse Survey Data: A Hierarchical Bayesian Modeling Framework to Account for Uncertainty." *Proceedings of the National Academy of Sciences* 117, no. 39: 24173–24179. https://doi.org/10.1073/pnas.1913050117.

Leyk, S., A. E. Gaughan, S. B. Adamo, et al. 2019. "The Spatial Allocation of Population: A Review of Large-Scale Gridded Population Data Products and Their Fitness for Use." *Earth System Science Data* 11, no. 3: 1385–1409. https://doi.org/10.5194/essd-11-1385-2019.

Liaw, A., and M. Wiener. 2002. "Classification and Regression by Randomforest." *R News* 2, no. 3: 18–22.

Marconcini, M., A. M.- Marconcini, T. Esch, and N. Gorelick. 2021. "Understanding Current Trends in Global Urbanisation—The World Settlement Footprint Suite." *GI_Forum 2021* 9: 33–38. https://doi.org/10.1553/giscience2021_01_s33.

Mesev, V., ed. 2003. "Zone-Based Estimation of Population and Housing Units From Satellite-generated Land Use/Land Cover Maps." In *Remotely-Sensed Cities*. CRC Press.

Metzger, N., J. E. Vargas-Muñoz, R. C. Daudt, et al. 2022. "Fine-Grained Population Mapping From Coarse Census Counts and Open Geodata." *Scientific Reports* 12, no. 1: 20085. Article 1. https://doi.org/10.1038/s41598-022-24495-w.

Muro, J., L. Zurita-Arthos, J. Jara, et al. 2020. "Earth Observation for Settlement Mapping of Amazonian Indigenous Populations to Support SDG7." *Resources* 9, no. 8: 97. Article 8. https://doi.org/10.3390/resources9080097.

Neal, I., S. Seth, G. Watmough, and M. S. Diallo. 2022. "Census-Independent Population Estimation Using Representation Learning." *Scientific Reports* 12, no. 1: 5185. Article 1. https://doi.org/10.1038/s41598-022-08935-1.

Ogrosky, C. E. 1975. "Population Estimates From Satellite Imagery." *Photogrammetric Engineering and Remote Sensing* 41, no. 6: 707–712.

OpenStreetMap contributors. 2023. Planet Dump. https://planet.osm.org.

Pesaresi, M., G. Huadong, X. Blaes, et al. 2013. "A Global Human Settlement Layer From Optical HR/VHR RS Data: Concept and First Results." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6, no. 5: 2102–2131.

Pesaresi, M., and P. Politis. 2022. *GHS-BUILT-S R2022A—GHS Built-Up Surface Grid, Derived From Sentinel2 Composite and Landsat, Multitemporal (1975–2030)*. https://doi.org/10.2905/D07D81B4-7680-4D28-B896-583745C27085.

Posit Team. 2023. RStudio: Integrated Development Environment for R. Posit Software, PBC. http://www.posit.co/.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Sanchez-Cespedes, L. M., D. R. Leasure, N. Tejedor-Garavito, et al. 2024. "Social Cartography and Satellite-Derived Building Coverage for Post-Census Population Estimates in Difficult-to-Access Regions of Colombia." *Population Studies* 78, no. 1: 3–20. https://doi.org/10.1080/00324728.2023.2190151.

Schiavina, M., M. Melchiorri, M. Pesaresi, et al. 2022. *GHSL Data Package 2022: Public Release GHS P2022*. Luxembourg: Publications Office of the European Union. Available at: https://human-settlement.emergency.copernicus.eu/documents/GHSL_Data_Package_2022.pdf.

Sirko, W., S. Kashubin, M. Ritter, et al. 2021. *Continental-Scale Building Detection From High Resolution Satellite Imagery* (No. arXiv:2107.12283). arXiv. https://doi.org/10.48550/arXiv.2107.12283.

Stan Development Team. 2023. *Stan Modeling Language Users Guide and Reference Manual* (Version 2.32) [Computer Software]. https://mc-stan.org.

Stevens, F. R., A. E. Gaughan, C. Linard, and A. J. Tatem. 2015. "Disaggregating Census Data for Population Mapping Using Random Forests With Remotely-Sensed and Ancillary Data." *PLoS One* 10, no. 2: e0107042. https://doi.org/10.1371/journal.pone.0107042.

United Nations. 2022. *Day of 8 Billion*. https://www.un.org/en/dayof8billion.

United Nations Department of Economic and Social Affairs, Population Division. 2024. *World Population Prospects 2024: Summary of Results* (No. UN DESA/POP/2024/TR/NO. 9; p. 80). https://desapublications.un.org/publications/world-population-prospects-2024-summary-results.

United Nations Statistics Division. (2023, December). *Census Dates*. Demographic and Social Statistics. https://unstats.un.org/unsd/demographic-social/census/censusdates/.

Walker, R. S., M. J. Hamilton, and A. A. Groth. 2014. "Remote Sensing and Conservation of Isolated Indigenous Villages in Amazonia." *Royal Society Open Science* 1, no. 3: 140246. https://doi.org/10.1098/rsos.140246.

Wardrop, N. A., W. C. Jochem, T. J. Bird, et al. 2018. "Spatially Disaggregated Population Estimates in the Absence of National Population and Housing Census Data." *Proceedings of the National Academy of Sciences* 115, no. 14: 3529–3537. https://doi.org/10.1073/pnas.1715305115.

Weber, E. M., V. Y. Seaman, R. N. Stewart, et al. 2018. "Census-Independent Population Mapping in Northern Nigeria." *Remote Sensing of Environment* 204: 786–798. https://doi.org/10.1016/j.rse.2017.09.024.

Weiss, D. J., A. Nelson, H. S. Gibson, et al. 2018. "A Global Map of Travel Time to Cities to Assess Inequalities in Accessibility in 2015." *Nature* 553, no. 7688: 333–336. https://doi.org/10.1038/nature25181.

WorldPop, & Institut National de la Statistique du Mali. 2022. *Census-Cartography-Based Gridded Population Estimates for Mali (2020), Version 1.0*. University of Southampton. https://doi.org/10.5258/SOTON/WP00745.

WorldPop, & National Population Commission of Nigeria. 2021. *Bottom-Up Gridded Population Estimates for Nigeria, Version 2.0*. University of Southampton. https://doi.org/10.5258/SOTON/WP00729.

Wu, S., X. Qiu, and L. Wang. 2005. "Population Estimation Methods in GIS and Remote Sensing: A Review." *GIScience & Remote Sensing* 42, no. 1: 80–96. https://doi.org/10.2747/1548-1603.42.1.80.