

Article

New Perspectives for Mapping Global Population Distribution Using World Settlement Footprint Products

Daniela Palacios-Lopez ^{1,*}, Felix Bachofer ¹, Thomas Esch ¹, Wieke Heldens ¹,
Andreas Hirner ¹, Mattia Marconcini ¹, Alessandro Sorichetta ², Julian Zeidler ¹,
Claudia Kuenzer ¹, Stefan Dech ¹, Andrew J. Tatem ² and Peter Reinartz ¹

¹ German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Oberpfaffenhofen, D-82234 Wessling, Germany; Felix.Bachofer@dlr.de (F.B.); Thomas.Esch@dlr.de (T.E.); Wieke.Heldens@dlr.de (W.H.); Andreas.Hirner@dlr.de (A.H.); Mattia.Marconcini@dlr.de (M.M.); Julian.Zeidler@dlr.de (J.Z.); Claudia.Kuenzer@dlr.de (C.K.); Stefan.Dech@dlr.de (S.D.); Peter.Reinartz@dlr.de (P.R.)

² WorldPop, Department Geography and Environment, University of Southampton, Southampton SO17 1B, UK; A.Sorichetta@soto.ac.uk (A.S.); A.J.Tatem@soton.ac.uk (A.J.T.)

* Correspondence: Daniela.PalaciosLopez@dlr.de; Tel.: +49-81-5328-2169

Received: 3 September 2019; Accepted: 28 October 2019; Published: 31 October 2019



Abstract: In the production of gridded population maps, remotely sensed, human settlement datasets rank among the most important geographical factors to estimate population densities and distributions at regional and global scales. Within this context, the German Aerospace Centre (DLR) has developed a new suite of global layers, which accurately describe the built-up environment and its characteristics at high spatial resolution: (i) the World Settlement Footprint 2015 layer (WSF-2015), a binary settlement mask; and (ii) the experimental World Settlement Footprint Density 2015 layer (WSF-2015-Density), representing the percentage of impervious surface. This research systematically compares the effectiveness of both layers for producing population distribution maps through a dasymetric mapping approach in nine low-, middle-, and highly urbanised countries. Results indicate that the WSF-2015-Density layer can produce population distribution maps with higher qualitative and quantitative accuracies in comparison to the already established binary approach, especially in those countries where a good percentage of building structures have been identified within the rural areas. Moreover, our results suggest that population distribution accuracies could substantially improve through the dynamic preselection of the input layers and the correct parameterisation of the Settlement Size Complexity (SSC) index.

Keywords: global population distribution mapping; World Settlement Footprint; percent impervious surface; dasymetric mapping; Settlement Size Complexity Index

1. Introduction

According to the latest revision of the United Nations (UN), World Population Prospects, the world's population is projected to grow from 7.7 billion in 2019 to 10.9 billion in 2100 [1]. Considered part of the four global demographic "megatrends", population growth next to population ageing, migration and urbanization, is an important indicator for economic, social and environmental development [2]. For this reason, accurate knowledge of the size, location, and distribution of the human population is fundamental for successfully achieving a sustainable future. An effective monitoring of global population change, allows implementing efficient government policies to allocate financial resources, plan interventions and quantify populations at risk.

To this end, since the late 1980s considerable efforts have been taken to produce global or continental scale, high resolution gridded population maps describing the spatial distribution of human population [3]. Over the last 20 years, the ongoing improvement in the availability and spatial detail of census population data, the better quality and spatial resolution of remote sensing data, the development of sophisticated geospatial analysis methods and the statistical refinement of modelling techniques, have been leveraged to produce more accurate datasets that capture the changes in magnitude, composition and distribution of human population over time [4].

Global or large scale gridded population datasets considered state-of-the-art in terms of open access archives of population distribution data include: the Rural–Urban Mapping Project (GRUMP) [5], the Gridded Population of the World, Version 4 (GPWv4) [6,7], the LandScan Global Population database [8,9], the Global Human Settlement Layer–Population grid (GHS-POP) [10,11], the WorldPop datasets [12–16] and the recently developed High Resolution Settlement Layer (HRSL) population grids [17]. Current and previous versions of these products have proved to be an important source of information and essential input for a wide range of cross-disciplinary applications including: poverty mapping [18–20], epidemiological modelling and disease burden estimation [21–23], interconnectivity and accessibility analyses [24–26], deriving past and future population estimates [15,27,28], disaster management [29–31] and human settlement characterisation [32] among others.

The modelling techniques of these population distribution datasets are based on a common methodology which consists of the disaggregation of census data from administrative units (polygons or source zones) into smaller areal units of fixed spatial resolution (grid cells or target zones) [3]. Population disaggregation is accomplished using two areal interpolation methods: areal-weighting and dasymetric mapping. With areal-weighting interpolation, a grid of fixed spatial resolution is intersected with the census polygons and each grid cell is assigned a portion of the total population based solely on the proportion of the area of the administrative unit that falls within each grid cell [9]. Thus far, the GPWv4 is the only dataset produced based on areal-weighting interpolation, while the rest of the population datasets employ a dasymetric mapping approach. This method seeks to improve the distribution of population through the incorporation of one or multiple geospatial covariates or categorical ancillary datasets that influence the variations in the densities and distribution of population within the administrative units [33].

The most commonly employed geospatial covariates include: land cover and land use types, intensity of nightlights, climatic factors, human settlements, urban/rural extents, water features, road networks and topographic elevation and slope. In this regard, LandScan and WorldPop population grids use multiple best-available local or global covariates that are statistically assessed to produce a weighted layer that is used as input in the dasymetric modelling method [8,12,16]. Here, the resulting population grids show an asymmetrical distribution of population counts per administrative unit, in which each grid cell is assigned a portion of the population depending on the individual calculated weights [34]. While sophisticated, this technique presents a number of limitations and disadvantages. For example, the assignment of relative weights to each individual covariate layer is subjective and based on local relationships [35,36]. In other words, the model is country-specific limiting the direct transferability of the model to global scales [8]. Moreover, temporal agreement between all covariate layers and population census data is difficult to achieve, restricting the production of a globally consistent population dataset. Finally, the use of multiple covariate layers reduces the applicability of the final population grids, as explained by Balk et al. [3].

In this framework, however, it has been demonstrated that not all the commonly used geospatial covariates are equally important for population disaggregation. According to the research presented by Nieves et al. [4], geographical data pertaining to the built-up environment and urban extents are the two most important covariates for predicting population densities and are significantly more important than other covariates at both regional and global scales. In this respect, the GHS-POP and HRSL population grids are processed using a binary-dasymetric mapping approach, restricting the distribution of population only to those grid cells identified as human settlements. The GHS-POP

uses the Global Human Settlement Layer built-up grids (GHSL-BUILT) [37], while the HRSL uses a binary mask of areas identified as human-made buildings extracted from very high-resolution satellite imagery. While this modelling approach is less complex and allows global transferability [33], the population mapping accuracies of these products largely depend on the complete identification of building structures and are affected by omission and commission errors [13].

In this context, the German Aerospace Center (DLR) has developed dedicated global layers and related analysis tools that describe the built-up environment and its characteristics with high accuracy and high spatial resolution. The first includes the Global Urban Footprint (GUF) dataset, which was released in 2016 [38]. The GUF was produced based on an operational framework that automatically processed and analysed over 180,000 TerraSAR-X/TanDEM-X radar images collected during 2011–2013. It provides a global human settlement map at 12 m resolution [39], which up to now, has been employed by more than 500 institutions for a broad scope of applications [40], including studies focused on population disaggregation [12,41,42]. Currently, the DLR is working on a suite of follow-on products—the World Settlement Footprint (WSF)—with an extended semantic depth, based on the joint analysis of Landsat 8 and Sentinel 1 optical and radar imagery [43]. The first two releases of this new suite will include: (i) a binary settlement mask named WSF-2015 outlining settlements globally at 10-m resolution; and (ii) the experimental WSF-2015-Density layer, which estimates the percent of impervious surface for the pixels labelled as settlement in the WSF-2015 [43].

Impervious surfaces are primarily associated with streets, sidewalks and building structures. They can be defined as surfaces consisting of materials such as asphalt, concrete or stone that seal the soil surface, eliminating water infiltration [44]. Impervious surfaces extracted using different remote sensing methodologies have been examined in a small number of population distribution studies [45,46]. In these studies, the authors have demonstrated that impervious surfaces are highly correlated to population counts, making them good predictors of population distribution. Nevertheless, these studies have only focused on limited areas, thus leading to results and methodologies that are not globally transferable. In the same way, in producing population distribution maps based on settlement extent products, Reed et al. [34] showed that an initial version of the WSF-2015 layer was capable of producing population distribution maps with predictive accuracies higher than the GHSL layer and relatively close to the HRSL layer, employing different population distribution methods. However, while currently the HRSL layer is available only for a limited number of countries, the novel WSF-2015 and the experimental WSF-2015-Density layers have the potential to become the ideal covariates to support population disaggregation methods and to produce global population distribution datasets with improved accuracy and higher spatial resolution than those currently available.

Following this premise, the main goal of this research was to examine the suitability of the WSF-2015 and the—thus far experimental—WSF-2015-Density layers as input covariates for the development of a new global population distribution dataset. Population distribution maps were produced using a dasymetric mapping approach in combination with the finest population census data available at global scale at the time of writing. We specifically focused on the systematic cross-comparison between the performance of the binary and the impervious layer, to investigate if quality and accuracy improvements in population disaggregation can be achieved with the WSF-2015-Density layer, compared to the already established binary approach that has been employed by other population datasets and their baseline settlement layers.

Through a comprehensive quantitative assessment, we evaluated the mapping performance of each covariate layer, addressing the influence of: (i) the spatial resolution of the input census data; (ii) the quality of the input covariate layers; and (iii) the spatial distribution of the built-up environment, on the final results.

The corresponding analyses were conducted for nine representative countries of different size and different levels of urbanisation and population aggregation.

2. Materials and Methods

2.1. Input Geospatial Covariates: WSF-2015 and WSF-2015-Density Layers

2.1.1. WSF-2015 Layer

The WSF-2015 is a novel layer outlining the extent of human settlements globally at 10-m resolution. The dataset has been derived by jointly exploiting multi-temporal Sentinel-1 Synthetic Aperture Radar and Landsat-8 optical satellite imagery collected during 2014–2015, of which ~107,000 and ~217,000 scenes have been processed, respectively.

The basic underlying hypothesis is that the dynamics of human settlements over time are sensibly different with respect to those of all other non-settlement classes. Accordingly, for all the scenes available for the given target region, key temporal statistics (e.g., temporal maximum, minimum, mean, variance, etc.) have been concurrently computed for: (i) the original Sentinel-1 backscattering; and (ii) different spectral indices extracted from the Landsat-8 data after removing clouds and cloud shadows. Next, candidate training samples for the settlement and non-settlement class have been extracted on the basis of specific thresholds determined—based on extensive empirical analysis—for some of the resulting temporal features. Classification is then performed separately for the optical- and radar-based features by means of Support Vector Machines (SVMs) and, finally, the two outputs are properly combined. The WSF-2015 exhibited high accuracy and reliability, outperforming all other existing similar global layers. Specifically, this has been quantitatively demonstrated through an extensive validation exercise performed in collaboration with Google where 900,000 reference samples have been labelled by crowd-sourcing photointerpretation for a collection of 50 globally distributed test sites of 1×1 lat/lon degree size. The layer is currently available for online browsing on the ESA Urban-TEP platform; furthermore, a comprehensive description of the classification system and the validation results is provided in [43].

2.1.2. WSF-2015-Density Layer

The WSF-2015-Density is one of the first experimental developments of the WSF product and service portfolio, aiming to enhance the semantic and thematic scope of the WSF-2015; in particular, the layer describes the percent impervious surface (PIS) within areas categorised as settlements in the WSF-2015. Effectively mapping the PIS is of high importance to assess—among others—the risk of urban floods, the urban heat island phenomenon as well as the reduction of ecological productivity. Furthermore, it is generally considered as an effective proxy for the housing density, thus making it particularly suitable for supporting spatial population distribution [45–47]. The current processing methodology follows the approach originally described by Marconcini et al. [48] and is based on the assumption that a strong inverse relation exists between vegetation and impervious surfaces (i.e., the higher the presence of vegetation is, the lower the corresponding imperviousness is). Accordingly, the core idea is to compute and analyse for each pixel the temporal maximum of the Normalised Difference Vegetation Index (NDVI), which depicts the status at the peak of the phenological cycle. To this purpose, the NDVI available from the TimeScan dataset [40,49] has been used, which has been derived globally from Landsat-8 scenes acquired during 2014–2015. Figure 1 shows a subset of the WSF-2015-Density layer for Toluca state in Mexico. Values range between 0 and 100, with red and green tones highlighting high and low PIS, respectively.

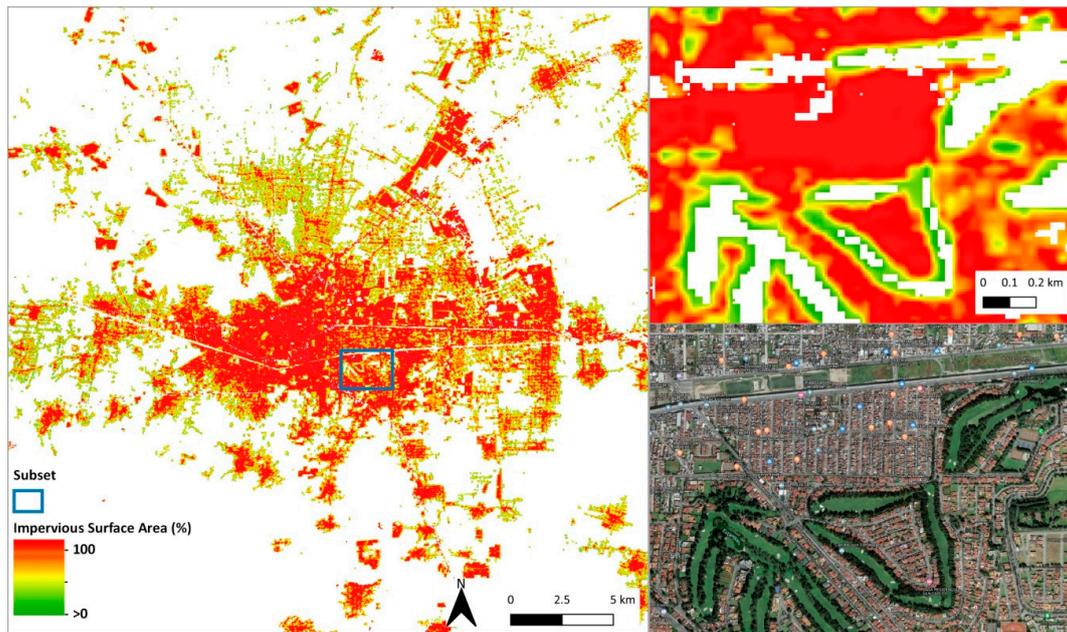


Figure 1. Covering the area of Toluca in Mexico and subset comparing the layer against VHR imagery. White areas: Pixels outside the WSF-2015 settlement mask.

2.2. Input Census Data

For this research, population census data for nine low-, middle- and highly urbanised countries [50] located in four different macro-regions of the world were collected to analyse how the differences in the level of spatial granularity of the available administrative boundaries and the variability in the morphology of built-up landscapes influence the accuracy of each covariate layer. To achieve these objectives, countries were selected on basis of the availability of population census data at different spatial aggregation levels. In other words, countries were selected only if the census data allowed for the spatial aggregation of the administrative boundaries up to four administrative levels.

The Center of International Earth Science Information Network (CIESIN) provided geographic administrative boundaries and corresponding population counts for Cambodia, Côte d'Ivoire, England, France, Germany, Malawi, Mexico, and Vietnam. CIESIN population data were selected for this research, as it has been used in the production of other population dataset such as GPWv4, GHS-POP, WorldPop and the HRSL. CIESIN collected census data at the highest spatial detail available from the results of the 2010 round of Population and Housing Censuses, which occurred between 2005 and 2014. CIESIN data include two types of population estimates: census-based and UN-adjusted, both estimated for the years 2000, 2005, 2010, 2015 and 2020. Initial population estimates were derived for each administrative unit by means of an exponential model fitted on at least two census counts for each country [17]. However, to allow for global comparisons, CIESIN adjusted the census counts to the target year of 2010, which were then then interpolated and extrapolated to produce the UN-adjusted estimates with the objective to correct for over- or under estimations [6,7]. The 2015 UN-adjusted estimates were used in this research.

For Myanmar, population data were collected from the Ministry of Immigration and Population in reference to the Population and Housing Census of 2014 [51] and was joined with publicly available geographic administrative boundaries [52]. The population data were released on May 2015 and the original population counts were used in this research.

For each country, administrative boundaries and population counts were aggregated at four levels of spatial resolution using attribute information stored within the data. Table 1 shows the total population for 2015 for each country as well as the official administrative unit nomenclature at each spatial aggregation level, the number of administrative units, the average area and the average spatial

resolution (ASR). The ASR is calculated as the square root of each country total area divided by the number of administrative units, representing the effective resolution units within each country [3].

Table 1. Input Census Data Characteristics.

Country (ISO)/Census Year	Total Population 2015	Official Admin. Unit Nomenclature	No. of Units	Average Area of Units (km ²)	ASR (km)
CIV Côte d'Ivoire 2014	22,701,552	Sub-Prefectures (Adm 3)	517	621.85	24.99
		Departments (Adm 2)	110	2907.6	54.17
		Region (Adm 1)	35	9220.92	96.03
		National (Adm 0)	1	322,744.29	568.11
DEU Germany 2014	80,688,539	Enumeration Area (EA Level)	11,292	31.26	5.59
		Districts (NUTS3)	402	878.25	29.64
		States (NUTS1)	16	22,066.28	148.55
		National (NUTS 0)	1	353,060.51	594.19
ENG England 2014	54,376,281	Enumeration Area (EA Level)	6791	19.2	4.38
		District (Adm 2)	326	400.16	20.00
		Region (Adm 1)	9	14,494.94	120.39
		National (Adm 0)	1	130,454.54	361.18
FRA France2009	64,395,348	Enumeration Area (EA Level)	36,562	15.09	3.89
		Departments (NUTS3)	96	5749.86	75.83
		Regions (NUTS2)	22	25093.51	158.41
		National (NUTS 0)	1	552,057.38	743.01
KHM Cambodia 2008	15,394,276	Commune (Adm 3)	1633	109.66	10.47
		District (Adm 2)	197	909.06	30.15
		Province (Adm 1)	25	7163.40	84.64
		National (Adm 0)	1	179,084.95	423.18
MEX Mexico 2010	129,731,190	Enumeration Area (EA Level)	65,477	27.7	4.91
		Municipality (Adm 2)	2456	804.65	25.36
		States (Adm 1)	32	59,898.45	222.15
		National (Adm 0)	1	1,579,248.33	1256.68
MMR Myanmar 2014	50,279,900	Township (Adm 3)	330	2032.66	45.09
		District (Adm 2)	74	9064.6	95.21
		Regions (Adm 1)	15	44,718.7	211.47
		National (Adm 0)	1	670,780.63	819.01
MWI Malawi 2010	17,215,235	Enumeration Area (EA Level)	12,550	7.19	2.68
		Traditional Authority (Adm 3)	357	252.92	15.90
		District (Adm 2)	32	2821.69	53.12
		National (Adm 0)	1	90,294.35	300.49
VNM Vietnam 2009	93,447,596	District (Adm 3)	688	477.52	21.85
		Municipality-Province (Adm 2)	63	5214.87	72.21
		Region (Adm 1)	6	54,756.19	234.00
		National (Adm 0)	1	328,537.15	573.18

2.3. Population Distribution: Dasymetric Mapping Approach

Population distribution maps for 2015 were generated for each country at each administrative unit level using a dasymetric mapping approach, where population census data from administrative boundaries (source zones) are disaggregated into smaller areal units of fixed spatial resolution (target zones). The size of the target zones is normally defined by the pixel resolution of the different ancillary datasets employed to restrict and refine the distribution of the population within each administrative unit [53]. The estimated population per grid cell is defined in Equation (1):

$$\text{Pop}_t = \text{Pop}_s \frac{A_t * W_p}{\sum_{t \in s} (A_t W_p)} \quad (1)$$

where Pop_t is the population of the target zone, Pop_s is the population of the source zone, A_t represents the area of the target zone and W_p is the weight of a grid cell within the target zone. With this modelling approach, population counts are maintained (volume-preserving property) at each original input source zone.

In this research, two types of dasymetric mapping techniques were used. The first method is the traditional binary approach, which relies on the WSF-2015 layer to assign a weighting factor of 1 to built-up pixels and a 0 for non-built-up pixels. The second method uses the WSF-2015-Density layer to assign a weighting factor that ranges from 0 to 100, estimating the PIS for the pixels classified as settlement in the WSF-2015.

2.4. Quantitative Accuracy Assessment

As stated by Bai et al. [54] “quantifying the accuracy of population distribution maps has been recognized as a critical and challenging task”. Determining the spatial and quantitative uncertainties of population distribution products is fundamental yet very difficult due to the lack of independent and compatible reference data [10]. Nevertheless, through well-established accuracy methods, it is possible to assess the effectiveness of new models (disaggregation methods and/or covariate layers) and investigate if higher population distribution accuracies can be reached in comparison to previous approaches. For this research, the accuracy of the two covariate layers was assessed by computing the difference between the estimated population counts extracted from maps produced using coarser administrative units (input units) and the actual population counts of the finest administrative units (validation units). This accuracy method has been widely employed in previous research [13,17,34,42,55]; however, it still presents some limitations, as high-resolution boundaries and population data (e.g., enumeration area level) are not publicly available for all countries.

For this reason, to gain a more comprehensive and detailed understanding of the mapping capabilities of each covariate layer, the final population distribution maps were evaluated following a series of thorough quantitative analyses performed at the validation unit level and the input level of the administrative units. The analysis at the validation unit level was divided in two parts. In the first part, an overall accuracy assessment was carried out to examine the influence of the spatial resolution of the input census data on the results. Here, population distribution maps were produced using three spatial aggregation levels of the administrative boundaries as input units (Analyses I–III in Table 2).

Table 2. Spatial aggregation levels of the administrative boundaries used as input units and validation units for each analysis (finest to coarser spatial detail) (EA, Enumeration Area).

Country (ISO)	Analysis	Level of Administrative Input Units	Level of Administrative Validation Units
KHM CIV MMR VNM	I	Adm 2	Adm 3
	II	Adm 1	
	III	Adm 0	
ENG	I	Adm 2	EA
	II	Adm 1	
	III	Adm 0	
FRA	I	NUTS 3	EA
	II	NUTS 2	
	III	NUTS 0	
DEU	I	NUTS 3	EA
	II	NUTS 1	
	III	NUTS 0	
MWI	I	Adm 3	EA
	II	Adm 2	
	III	Adm 0	
MEX	I	Adm 2	EA
	II	Adm 1	
	III	Adm 0	

For each analysis, four main descriptive statistics were calculated to measure the overall accuracies of each layer. These metrics are briefly described in Table 3 and include: the MAE (Mean Absolute Error), the MAPE (Mean Absolute Percentage Error), the Root Mean Square Error (RMSE) and the coefficient of determination (R^2).

Table 3. Descriptive statistics for overall accuracy assessment at the validation unit level for Analyses I–III.

Metric	Description
$MAE_i = \frac{\sum_{VU=1}^n PE_{VU} - P_{VU} }{n} \quad (2)$	MAE is the mean absolute error at each level of analysis (i), calculated as the average of the sum of the absolute differences between the estimated population (PE_{VU}) and the actual population (P_{VU}) at each validation unit.
$MAPE_i = \frac{MAE_i}{Av. Pop} \times 100\% \quad (3)$	MAPE is the mean absolute percentage error at each level of analysis (i), calculated as the MAE_i divided by the average population of each country.
$RMSE_i = \sqrt{\frac{\sum_{VU=1}^n (P_{VU} - PE_{VU})^2}{n}} \quad (4)$	RMSE is the root mean square error at each level of analysis (i), calculated as the square root of the mean of the sum of squares of the differences between the estimated population at (PE_{VU}) and the actual population (P_{VU}) at each validation unit.
R^2	Defined as the coefficient of determination at each level of analysis, derived from classical linear least square modelling with constant intercept at 0. It is also defined as the square of the Pearson correlation coefficient, to measure the variation between the estimated population and the actual population of all validation units. Readers can refer to [56] for detailed calculations.

The second part of the analysis was carried out only for the population maps produced using the finest input units (Analysis I in Table 2). Here, similar to the methodology and classification presented by Bai et al. [54], the Relative Estimation Error (REE) metric was used to identify the amount and distribution of error produced by each covariate layer. The REE for each validation unit was calculated as:

$$REE_{VU} = ((PE_{VU} - P_{VU}) / P_{VU}) * 100\% \quad (5)$$

where PE_{VU} is the estimated population of the validation unit and P_{VU} is the actual population of the validation unit. Using the REE_{VU} , validation units were grouped and classified into different REE ranges (Table 4).

Table 4. REE classification [54].

REE Ranges	Description
[-100%, -50%)	Greatly underestimated
[-50%, -25%)	Underestimated
[-25%, 25%]	Accurately estimated
(25%, 50%]	Overestimated
(50%, ≥100%]	Greatly overestimated

From this classification, two sub-analyses were conducted for each country. First, for a better understanding of the error distribution associated with each covariate layer, we calculated the percentage of each country's total population that fell within each error range. Second, for each country, we calculated the average actual population and average number of settlements pixels for the validation units that fell within each error range. This last analysis was done to identify if there is any

relationship between the amount of population that needs to be distributed (P_{VU}) and the number of available settlement pixels, and if the ratio between these two parameters can explain the REE values reported in the validation units.

Finally, as the reported accuracy at the validation unit level is only a reflection of the capability of each input covariate layer to correctly allocate population counts at the input unit level, a series of analyses were carried out at the input unit level, focusing only on Analysis I (Table 2). First, we used the RMSE metric as a summary of the error within each original input unit, following the methodology presented by Mennis and Hultgren [57]. RMSE was calculated as the square root of the mean of the sum of squares of the difference between the actual population counts and estimated population counts of all validation units within an input unit:

$$RMSE_{IU} = \sqrt{\frac{\sum_{VU \in IU} (P_{VU} - PE_{VU})^2}{n}} \quad (6)$$

where P_{VU} is the actual population at validation unit, PE_{VU} is the estimated population at validation unit and n is the number of validation units within an input unit. To compare the effectiveness of the covariate layers, input units were grouped according to the layer that produced the lowest RMSE values and for each group the percentage of each country's total population was calculated.

Second, on basis of these results, we undertook a series of analyses to identify and describe the regions where one layer outperformed the other. For this analysis, we derived the Settlement Size Coverage (SSC) index, which classifies each input unit according to (i) the number of small-, medium- and large-settlement objects that can be found within each unit; and (ii) the proportion of each input unit's total area that is covered by these settlements objects. To calculate the SSC index, settlement objects were created, where each object is composed of connected settlement pixels via at least one pixel edge or corner (8-neighbourhood), as described by Esch et al. [58]. The SSC index within each given input unit was derived as:

$$SSC_{IU} = \left(\frac{\# \text{settlement pixels}}{\# \text{settl. objects}} \right) * \left(\frac{\text{Sum of the area settl. objects}}{\text{Total area of input unit}} \right) * \left(\frac{\text{Area of largest settl. objects}}{\text{Mean area of settl. objects}} \right) \quad (7)$$

where high SSC_{IU} values indicate dense built-up environments and low SSC_{IU} values indicate sparse built-up environments. To allow country cross-comparisons, we normalised the SSC index values from 0 to 10 and divided it into three classes, as shown in Table 5. Thresholds were visually derived and evaluated against all available countries. For each SSC class, we calculated the average RMSE produced by each layer.

Table 5. Settlement Size Complexity Index classification scheme.

SSC Index Class	Description
Low (>0–1)	Small size settlements and low coverage of the total area of the input units
Medium [1–1.8)	Mix of small and medium size settlements and medium coverage of the total area of the input units
High [1.8–10)	Mix of medium and large size settlements with high coverage of the total area of the input units

3. Results

3.1. Visual Assessment of the Population Distribution Maps

The WSF-2015-Density and the WSF-2015 layers were used to produce population distribution maps for each country at each spatial aggregation level of the administrative units, representing the estimated night-time population (population counted at place of domicile) as the number of people

per grid cell for the year 2015. The final spatial resolution of the population distribution maps equals the spatial resolution of the input covariates (~10-m at the equator).

Because the volume of results (72 population distribution maps) is too large to present here in full, we focused on one representative country to visually inspect the thematic differences between the maps produced using the WSF-2015-Density and the WSF-2015 layers before turning to the quantitative analyses of all the maps. Figure 2 shows the final population distribution maps produced using the finest administrative units for Germany (enumeration areas), depicting the local metropolitan areas of Berlin and Munich. Note that, for the finest administrative units, these two areas have been modelled using a single census unit were local differences between the binary and the weighted disaggregation approaches are rather clear.

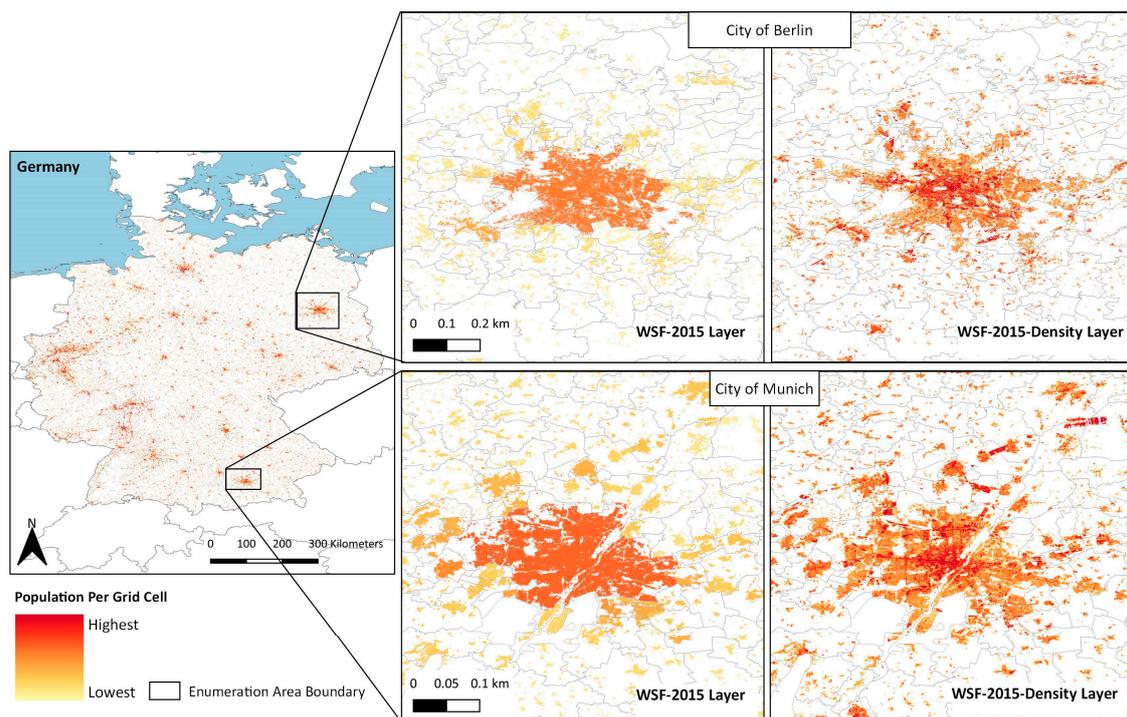


Figure 2. In this example: Estimated population as the number of people per grid cell for Germany in 2015 produced at the finest aggregation level of the input data (enumeration areas). The population distribution is displayed as the result of dasymetric approach using the WSF-2015 layer and the WSF-2015-Density layer. Detailed examples show the metropolitan areas of Berlin and Munich.

As illustrated in Figure 2, population disaggregation based on the WSF-2015 layer produces homogeneous population counts within each administrative unit in comparison to the WSF-2015-Density layer, which offers more spatial heterogeneity. As a result of the proportional allocation produced by the binary layer, it is possible to observe abrupt changes from high to low population counts between neighbouring administrative units. The transitions are considerably smoother when using the WSF-2015-Density layer, due to the weight given by the percent of impervious surface, which rarely changes abruptly at the boundaries of the administrative units.

3.2. Accuracy Assessment

3.2.1. Analyses at the Validation Unit Level

A summary of the accuracy assessment results using the WSF-2015-Density and WSF-2015 layers is presented in Table 6. Results show that, for each layer and each country, the highest R^2 , the lowest MAE, the lowest MAPE and the lowest RMSE values are reached using the finest administrative input

units (Analysis I, Table 2). Furthermore, from one level of spatial aggregation to the next, the values for the R^2 decrease, while the MAE, MAPE and RMSE values increase.

From the RMSE and MAE metrics, it can be seen that, for Analysis I, for most countries, errors remain below the size of the average population using any of the two covariate layers. While for all countries the MAE values remain below the average population size for Analyses I–III, RMSE exceeds this threshold in Analysis II in Germany and France and in Analysis III in Mexico and Myanmar. Additionally, the difference between the RMSE and the MAE values tends to increase as the spatial detail of the input units decreases, with significant higher differences in countries such as Côte d’Ivoire, France, Myanmar and Vietnam. In case of the three latter, the large differences can be explained by the large variances between the errors of the validation units within each country.

Comparing the results between the WSF-2015-Density and the WSF-2015 layers, it can be seen that, for Cambodia and Malawi, the best overall accuracies are reported using the WSF-2015 layer at all levels of aggregation. For the rest of the countries, the WSF-2015-Density layer performs better at all levels of aggregation, except for Mexico and Myanmar where there is a transition between layers in Analysis III.

Focusing only on the population distribution maps produced using the finest input units (Analysis I, Table 2), further analyses were performed at the validation unit level. First, to understand the amount of error produced by each input covariate layer, we calculated the percentage of each country’s total population that fell within each REE range. Classifying the REE values in different error ranges (Table 4), we calculated the percentage of each country’s total population that fell within each REE range for each covariate layer, as shown in Figure 3 and Table 7.

The percentage bar charts in Figure 3 show that, for each country, both covariate layers distribute approximately the same amount of population with comparable accuracies. From here, it can be seen that, for all countries, the largest percentage of the population was “accurately estimated” with estimation errors ranging from -25% to 25% for both covariate layers. For Côte d’Ivoire, Germany, England and Myanmar, this represents more than 50% of the total population; for France, Cambodia and Vietnam, between 40% and 50% of the total population; and, for Malawi and Mexico, between 30% and 40% of the total population. Moreover, for the majority of the countries, the second largest percentage of the population was either “underestimated” or “overestimated” (from ± 25 to $\pm 50\%$). For all countries, less than 15% of the total population was underestimated, while, for most countries, except Germany and Myanmar, from 15% to 25% of the total population was overestimated. Finally, the smallest percentage of the population for all countries was “greatly underestimated” or “greatly overestimated” ($\geq 50\%$ or $\leq -50\%$), with Malawi reporting an average of $\sim 30\%$ of the total population within these ranges, followed by Mexico with $\sim 25\%$, and France and Vietnam with $\sim 17\%$.

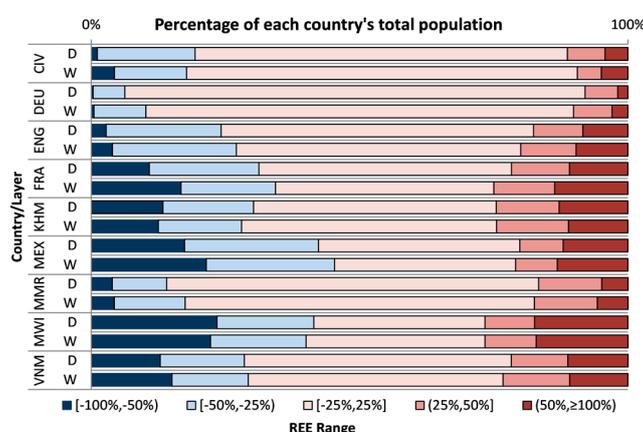


Figure 3. Percentage of each country’s total population that fell within each REE range. D, using the WSF-2015-Density layer; W, using the WSF-2015 layer.

Table 6. Accuracy assessment results using the WSF-2015 and the WSF-2015-Density covariate layers. Values of MAE and RMSE represent number of people.

Country ISO	Average Population	Analysis	No. of Input Unit	No. of Validation Units	WSF-2015-Density				WSF-2015			
					MAE	MAPE (%)	RMSE	R ²	MAE	MAPE (%)	RMSE	R ²
CIV	43,910.16	I	110	517	10,029.04	22.84%	40,198.00	0.7803	10,375.16	23.63%	44,814.26	0.7224
		II	35		11,851.45	26.99%	41,343.98	0.7725	11,862.96	27.02%	45,593.19	0.7130
		III	1		15,016.82	34.20%	47,045.80	0.5684	15,118.44	34.43%	50,124.64	0.3891
DEU	7145.64	I	402	11,291	828.86	11.60%	2261.67	0.9975	984.10	13.77%	2824.88	0.9961
		II	16		1897.35	26.55%	12,580.46	0.9316	2281.22	31.92%	14,409.45	0.9094
		III	1		2481.30	34.72%	23,280.14	0.9170	2999.64	41.98%	26,407.33	0.9010
EN	8007.11	I	326	6791	2218.00	27.70%	3309.71	0.1744	2347.93	29.32%	3401.02	0.1415
		II	9		2776.75	34.68%	4310.88	0.1000	3208.51	40.07%	4619.30	0.0474
		III	1		3098.81	38.70%	4666.95	0.0634	3642.90	45.50%	5017.18	0.0167
FRA	1761.26	I	96	36,562	589.00	33.44%	4605.53	0.8777	685.47	38.92%	5242.17	0.8352
		II	22		702.31	39.88%	9543.18	0.7698	817.24	46.40%	10,950.33	0.6333
		III	1		821.41	46.64%	11435.96	0.5279	954.06	54.17%	12495.90	0.3390
KHM	9426.99	I	197	1633	3425.38	36.34%	4898.26	0.6174	3241.26	34.38%	4694.16	0.6204
		II	25		4325.54	45.88%	6680.15	0.5244	4078.17	43.26%	6027.73	0.5371
		III	1		4738.49	50.27%	8363.82	0.5333	4343.88	46.08%	6270.24	0.5662
MEX	2915.00	I	2456	65,477	954.40	32.74%	2424.57	0.3841	1031.51	35.39%	2599.99	0.3672
		II	32		1080.44	37.06%	2440.33	0.3176	1194.89	40.99%	2611.97	0.3162
		III	1		1719.04	58.97%	30507.37	0.2326	1702.60	58.41%	3464.93	0.2604
MMR	76,263.92	I	75	330	32,257.60	42.30%	47,374.91	0.8214	34,301.82	44.98%	49,602.98	0.7986
		II	15		41,755.91	54.75%	58,807.41	0.7611	44,506.83	58.36%	64,708.38	0.7071
		III	1		83,960.45	110.09%	111,546.15	0.5243	66,606.76	87.34%	88,449.93	0.4051
MWI	1371.73	I	357	12,550	712.08	51.91%	1038.03	0.3231	687.40	50.11%	1001.41	0.3290
		II	32		795.36	57.98%	1219.17	0.1732	766.46	55.88%	1177.45	0.2050
		III	1		836.53	60.98%	1310.94	0.1924	792.69	57.79%	1182.53	0.2423
VNM	135,824.99	I	63	688	46,646.67	34.34%	76,804.15	0.6018	47,837.20	35.22%	87,481.13	0.5218
		II	6		57,187.23	42.10%	94,536.29	0.4317	61,288.84	45.12%	99,151.92	0.3578
		III	1		61,323.29	45.15%	95,472.76	0.3617	63,825.03	46.99%	100,829.93	0.2636

To identify if there is any significant relationship between the actual population to distribute in a particular validation unit and the number of available settlement pixels, we calculated the average actual population and the average number of settlement pixels for the validation units that fell within each REE range. Figure 4a shows the ratio between these two parameters for each REE range, where the general tendency indicates that, for most countries, errors of underestimation are mainly reported in validation units where a relatively low number of settlement pixels were identified in comparison to the average actual population reported for those validation units. In other words, errors of underestimation tend to increase as the ratio between the population and the number of settlement pixels increases. On the other hand, for most countries, errors of overestimation tend to increase as the ratio between the average actual population and the number of settlement pixels decreases, indicating that a large number of settlement pixels have been detected in relation to the average actual population reported for those validations units.

Table 7. Summary of the percentage of each country's total population that fell within each REE range.

	REE Range	[-100%, -50%)		[-50%, -25%)		[-25%, 25%]		(25%, 50%]		(50%, ≥100%]	
		D	W	D	W	D	W	D	W	D	W
CIV	%Population	1.11%	4.30%	18.22%	13.44%	69.40%	72.84%	6.98%	4.47%	4.29%	4.95%
DEU	%Population	0.34%	0.51%	5.90%	9.63%	85.78%	79.69%	6.05%	7.19%	1.92%	2.97%
ENG	%Population	2.78%	3.94%	21.40%	23.09%	58.22%	53.00%	9.21%	10.30%	8.39%	9.67%
FRA	%Population	10.82%	16.73%	20.42%	17.57%	47.06%	40.70%	10.79%	11.34%	10.92%	13.66%
KHM	%Population	13.35%	12.50%	16.87%	15.48%	45.23%	47.55%	11.73%	13.40%	12.82%	11.07%
MEX	%Population	17.37%	21.42%	24.97%	23.90%	37.50%	33.73%	8.09%	7.76%	12.07%	13.19%
MMR	%Population	3.92%	4.27%	10.14%	13.22%	69.30%	65.06%	11.74%	11.73%	4.92%	5.73%
MWI	%Population	23.44%	22.23%	18.03%	17.80%	31.87%	33.33%	9.25%	9.54%	17.41%	17.11%
VNM	%Population	12.84%	15.04%	15.66%	14.20%	49.78%	47.47%	10.50%	12.43%	11.23%	10.86%

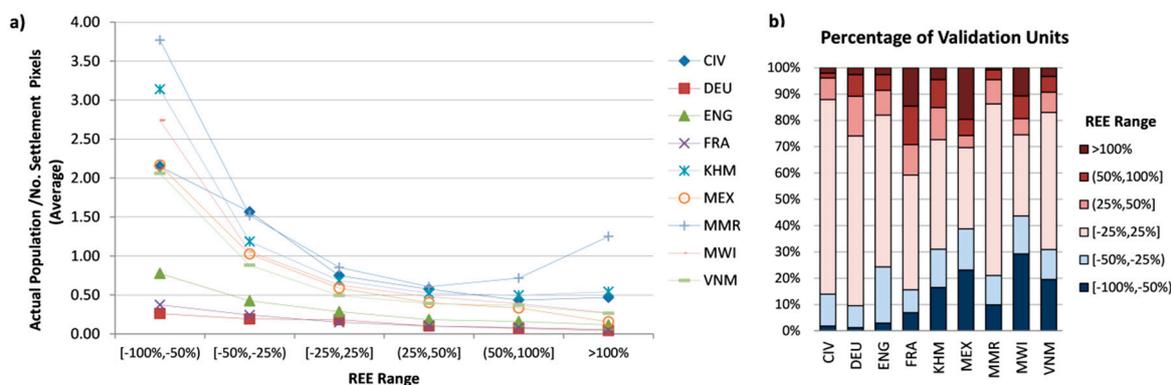


Figure 4. REE distribution: (a) ratio between the average population and the average number of settlement pixels for the validation units that fell within each REE range; and (b) percentage of validation units that fell within each REE range.

For a better understanding of the error distribution, the percentage of validation units that reported similar ratios and fell within each REE was quantified for each country. From the percentage bar charts in Figure 4b, it is possible to observe that, for countries such as Cambodia, Mexico, Malawi and Vietnam, more than 30% of the validations units reported errors of underestimation (from -100% to -25%), with Mexico, Malawi and Vietnam reporting $\sim 20\%$ of the validation units “greatly underestimated” (from -100% to -50%). In the same way, France reported the largest percentage of the validation units ($\sim 41\%$) with errors of overestimation (from 25% to $\geq 100\%$), followed by Mexico ($\sim 30\%$), Malawi and Germany ($\sim 25\%$). Here, Mexico reported the largest percentage of validation units “greatly overestimated”, with $\sim 20\%$ of the validation units with REE larger than 100% .

3.2.2. Analyses at the Input Unit Level

To evaluate the actual performance of each covariate layer, results at the validation unit level were used to calculate the $RMSE_{IU}$ metric of the original input census units used for population disaggregation according to Equation (6). Input units were grouped according to the input covariate layer that produced the lowest $RMSE_{IU}$ values and for each group the percentage of each country's total population was calculated.

Figure 5 illustrates the percentage bar charts for each country. As one can notice, for Germany, France and Mexico, the predominance of the WSF-2015-Density is clear, distributing more than 75% of each country's total population with overall lower RMSE values in comparison to the WSF-2015 layer. On the other hand, for Cambodia and Malawi, the WSF-2015 layer performs better, distributing more than 75% of the population more accurately compared to WSF-2015-Density layer. In the rest of the countries (i.e., Côte d' Ivoire, England, Myanmar and Vietnam), both layers perform equally, with the WSF-2015-Density layer distributing a slightly larger amount of the population better than the WSF-2015 layer.

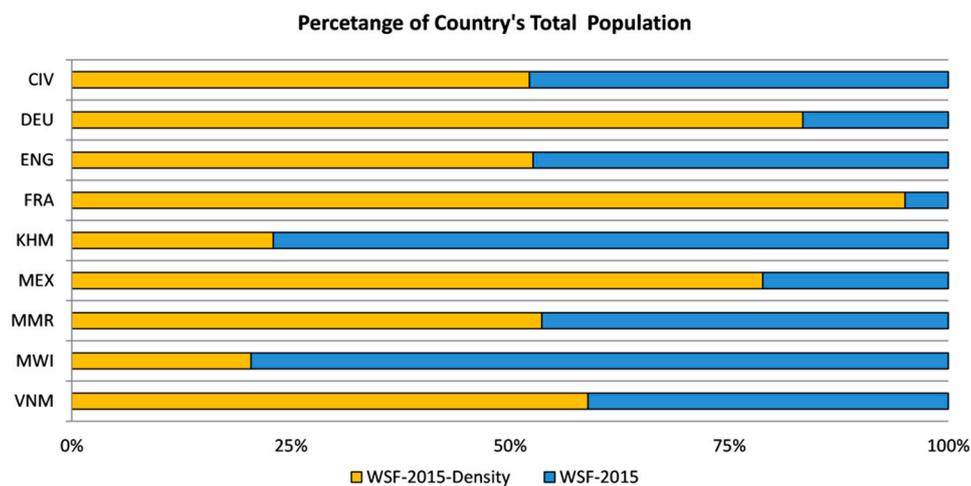


Figure 5. Percentage bar-charts of each country's total population distributed with higher accuracy by each covariate layer. Orange bars, WSF-2015-Density layer; Blue bars, WSF-2015 layer.

To identify the regions where each covariate layer produced higher accuracies, the input units of each country were classified according to the SSC index (Equation (7), Table 5). The map in Figure 6 illustrates the results of this classification for Côte d' Ivoire. Here, most of the input units fell within the "low" SSC class, which is characterised by small size settlement objects that cover a low percentage of each input unit's total area. A few input units fell within the "medium" SSC class, characterised by a mix of medium and small size settlements objects, and only two input units fell within the "high" SSC class, characterised by large size settlement objects that cover a large extent of each input unit's total area. For Côte d' Ivoire, some of the most populated cities are located within the "high" and "medium" input units, such as Abidjan, Bouake, Korhogo and Divo.

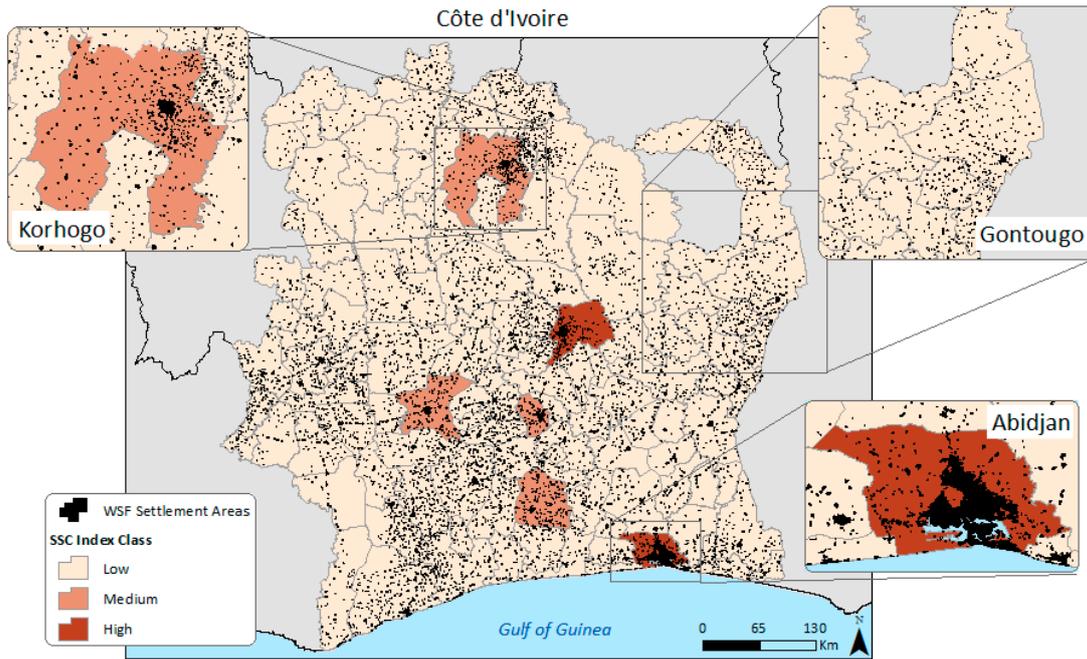


Figure 6. Input census units classified according to the SSC-Index for Côte d'Ivoire.

Following this classification, the same analysis was carried out for each country. Figure 7 shows the percentage of each country's total area (pie charts) and corresponding population (boxes) derived from the input units according to the SSC index.

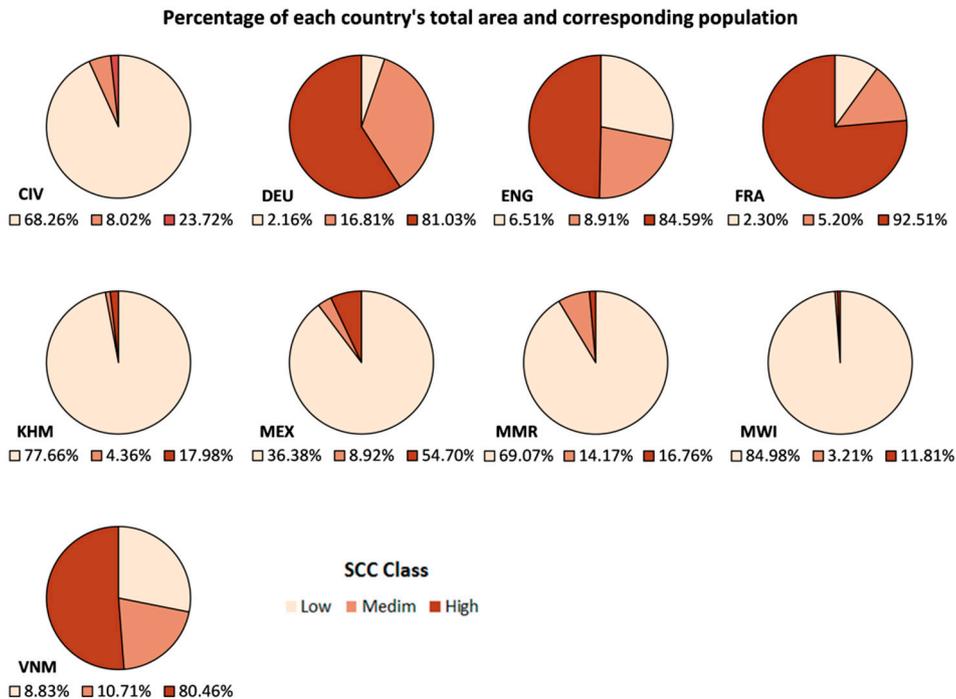


Figure 7. Percentage of each country's total area (pie charts) and corresponding population (boxes), classified according to the SSC index.

For Côte d' Ivoire, Cambodia, Mexico, Myanmar and Malawi, the largest percentage of the total area fell within the "low" SSC class. For all these countries, more than 70% of the population is located within these areas, except for Mexico, where the majority of the population (54.79%) is located within

areas belonging to the “high” SSC class. For Germany, England, France and Vietnam, the largest percentage of the total area fell within the “high” SSC class, where more than 80% of the population is located. For most countries, the second largest percentage of the area fell within the “medium” SSC class. In these areas, the second largest percentage of the population is located, however it does not exceed more than 17% of the total population.

For each SSC class, we computed the average RMSE error produced by each covariate layer and the percentage difference between the two layers (Table 8). The results indicate that, for all countries, the WSF-2015-Density layer performed better in regions that fell within the “high” SSC class, with improvements ranging from 1.12% to 31.20% over the WSF-2015 layer. For regions within the “low” or “medium” SSC classes, the behaviour of the covariate layers is more variable among the countries. For Germany, England, France and Myanmar, the WSF-2015-Density layer performed better for regions within the “low” SSC class, with improvements ranging from 4.36% to 22.40%, while, for Côte d’Ivoire, Cambodia, Malawi and Vietnam, the WSF-2015 layer performed better with improvements ranging from 2.12% to 9.82%. For the regions within the “medium” SSC class, the WSF-2015-Density layer performed better in Germany, England, France, Malawi and Vietnam, with improvements ranging from 6.62% to 21.03%, as opposed to Côte d’Ivoire, Cambodia, Mexico and Myanmar, where the WSF-2015 layer performed better with improvements ranging from 6.69% to 30%.

Table 8. RMSE (number of people) and percentage difference reported for each covariate layer at each SSC index class. D, results of the WSF-2015-Density layer; W, results of the WSF-2015 layer; positive bold values, countries where the WSF-2015-Density performed better; negative values, countries where the WSF-2015 performed better.

	Low SSC Class			Medium SSC Class			High SSC Class		
	RMSE (D)	RMSE (W)	%Diff.	RMSE (D)	RMSE (W)	%Diff.	RMSE (D)	RMSE (W)	%Diff.
CIV	6195.88	5824.85	−6.17%	13,385.74	9893.41	−30.00%	121,500.76	138,430.55	+13.03%
DEU	598.65	701.99	+15.89%	1169.94	1422.94	+19.51%	1715.78	2100.37	+20.16%
ENG	2449.15	2879.85	+16.16%	2580.89	3013.39	+15.46%	2908.04	2980.60	+2.46%
FRA	517.12	647.56	+22.40%	975.40	1207.01	+21.03%	4391.66	5124.74	+15.41%
KHM	4041.02	3785.02	−6.54%	3536.39	3084.83	−13.64%	6372.06	6443.97	+1.12%
MEX	892.80	874.05	−2.12%	2107.69	2253.53	−6.69%	2376.74	2626.13	+9.97%
MMR	33,452.74	34,943.76	+4.36%	39,432.66	32,580.79	−19.03%	43,682.69	59,832.04	+31.20%
MWI	819.79	768.93	−6.40%	778.43	831.73	+6.62%	1150.90	12,20.03	+5.83%
VNM	47,476.56	43,030.73	−9.82%	32,471.05	27,000.96	+18.40%	63,679.29	65,272.30	+2.47%

4. Discussion

In the above sections, we present a set of comprehensive analyses to compare the relative accuracies of population distribution maps produced using the WSF-2015 and the experimental WSF-2015-Density layers. The first analysis consisted of an overall accuracy assessment carried at the validation unit level, where metrics such as MAE, MAPE, RMSE and R^2 (Table 3) were used to evaluate maps produced using three spatial aggregation levels of the administrative units (Table 2). The results presented in Table 6 show that, for all countries and both covariate layers, the highest accuracy values were reported for population maps produced using the finest input census units (Analysis I, Table 2), with accuracies decreasing from one level of spatial aggregation to the next. These results are directly in line with previous findings [17,35,55], and confirm the premise that higher accuracies in population mapping can be achieved with improvements in the resolution of the input census data. In the same way, from a comparative point of view, the overall accuracy results showed that, for the majority of the countries, except Cambodia and Malawi, the WSF-2015-Density layer performed better than the WSF-2015.

When interpreting and comparing the overall accuracy results between countries and between covariate layers, there are, however, a set of considerations that need to be taken into account. First, it is important to understand, that regardless of the input covariate layer used for population disaggregation, high accuracies can be reached, when the number and ASR of the of the administrative units used for validation are similar to those of the administrative units used as input data (Table 2). This can be seen, for example, by examining the results of Analysis I for Côte d’Ivoire, Myanmar and Vietnam (Table 6).

The fact that these countries reported relatively good accuracy results is more likely to be due to the small difference between the number of administrative units used as input and validation units (407, 225 and 625, respectively) and the small ratio between their ASR (2.16, 2.11 and 3.30, respectively). These results are linked to the scale effect of the modifiable areal unit problem (MAUP), where the correlation between variables increases as the areal unit size becomes similar [59].

A second consideration to keep in mind is to avoid the use of the R^2 metric as unique statistical indicator to report the accuracy of population distribution models. Previous research has demonstrated that the lack of variability in the data influences the coefficient of determination [60]. For example, for England, where significantly low R^2 values were obtained in comparison to the MAE, MAPE and RMSE metrics, these can be related to the fact that the original census data reports similar population counts for a large number of the administrative units used for validation. This can be seen in the boxplots of Figure 8, where the reported actual population counts of the validation units of England are constrained within a small range of values. This small variability in the data, according to Goodwin et al. [60], results in a poor correlation between the estimated population counts and actual population counts as exemplified in the scatter plots of Figure 9. Here, it is possible to observe an amorphous or non-structured appearance of the data points for England in comparison to France, which results in a poor correlation, signalled by the almost horizontal trend-line.

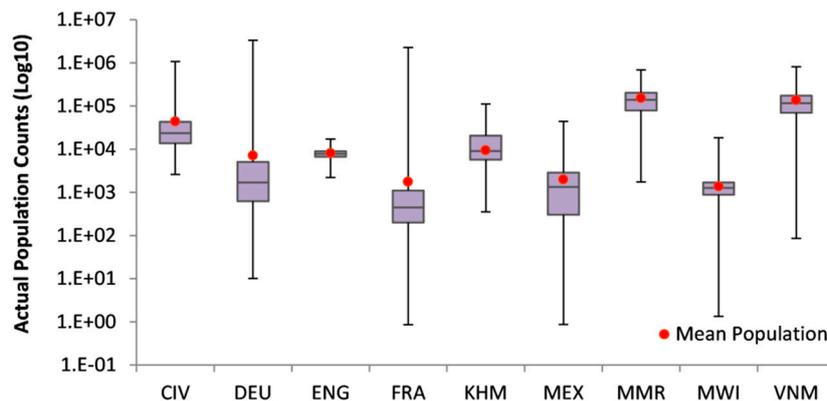


Figure 8. Boxplots of the distribution of the actual population counts of the validation units for each country with the inter-quartile range demarcated by the purple box.

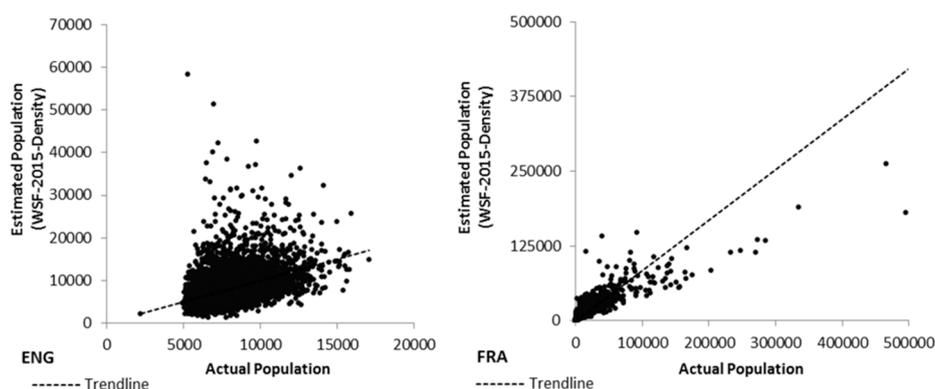


Figure 9. Scatter plot of estimated population and actual population for England and France at the validation unit level. Data show the results of population estimates using the WSF-2015-Density layer.

The aforementioned findings indicate that the use of single statistics metrics can be misleading and that population distribution maps can report high accuracy results independently of the quality of the underlying covariate layers used for population disaggregation. Therefore, it is important to emphasise, not only that full dissemination of the data used for modelling and validation is essential

when reporting accuracy results [3,6,35,54], but also that, to evaluate the real effectiveness of the covariate layers, it is necessary to undertake more in-depth analyses using complementary metrics.

In this research, with the use of the REE statistical metric (Equation (5)), it was possible to evaluate the amount and distribution of error generated by each covariate layer (Figure 3 and Table 7), and identify the areas where large errors of underestimation and large errors overestimation can be expected (Figure 4). Our results show that both layers perform similarly, distributing approximately the same percentage of each country's total population with the same REE values. For all countries, the largest percentage of the population has been estimated with errors ranging from -25% to 25% , which in previous research has been considered as "accurately estimated" [54]. Nevertheless, only in Côte d'Ivoire, Germany, England and Myanmar this represent more than 50% of the total population, which indicates that, for the rest of the countries, a significant percentage of the total population was distributed with larger errors of underestimation and errors overestimation.

We attribute these errors to the quality (completeness) of the covariate layers and to the fact that they do not take into account information on the land or building use. On the one hand, our findings indicate that errors of underestimation are reported in validation units where not enough settlement pixels have been found for population disaggregation. These errors increase as the ratio between the actual population and the number of settlement pixels increases (Figure 4a). This means, for example, that, in countries where a large percentage of the population and validation units were "greatly underestimated" (Table 7 and Figure 4b), such as France, Cambodia, Mexico and Malawi, this can be explained by the large amount of validation units where zero or very few settlement pixels have been identified (Figure 10). Therefore, despite the fact that the thematic accuracy of the WSF-2015 layer clearly outperforms any of the currently existing global human settlements masks [43], it is clear the data still show limitations with respect to a complete detection of all building structures. This can be explained by the spatial resolution of the Sentinel-1 and Landsat imagery used as input data, which restricts the identification of building structures, especially in regions where the settlement pattern is characterised by wide-spread single houses or very small hamlets.

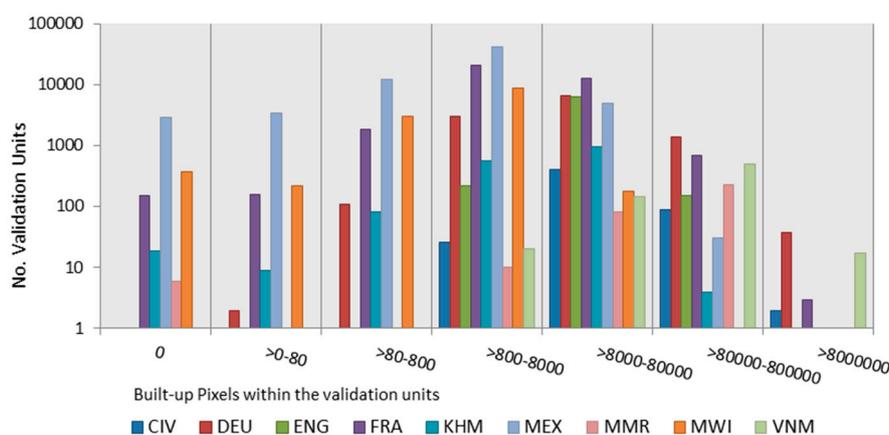


Figure 10. Number of settlement pixels identified within the validation units.

On the other hand, errors of overestimation are reported in validation units where a large number of settlement pixels have been reported in comparison to the amount of actual population, and that they increase as the ratio between these two parameters decreases (Figure 4a). After a visual analysis of VHR satellite imagery, we found that large errors of overestimation are mainly reported in validation units where seaports and industrial complexes exist. Figure 11 shows an example of the population distribution results for an input unit in England with this particular built-up environment. The red line represents the geographical boundary of the input unit used for population disaggregation and the blue lines represent the geographical boundaries of the validation units. Here, it is possible to observe industrial areas in the southern parts of the input unit. These areas capture a large amount of the population counts comparable to high-density residential areas, reporting large errors of overestimation

in the validation units. In the selected validation unit, for example the WSF-2015-Density layer reported a higher REE (186.56%) in comparison to the WSF-2015 (154.49%). This does not mean, however, that in every validation unit where this built-up environment exists the binary layer will perform better than the impervious layer. Depending on the extent and geographical boundaries of the input units, industrial or port areas can be mixed with residential areas, influencing the performance of each layer. More detailed information on and discussion of this aspect is provided at the end of this section in the context of the SSC index.

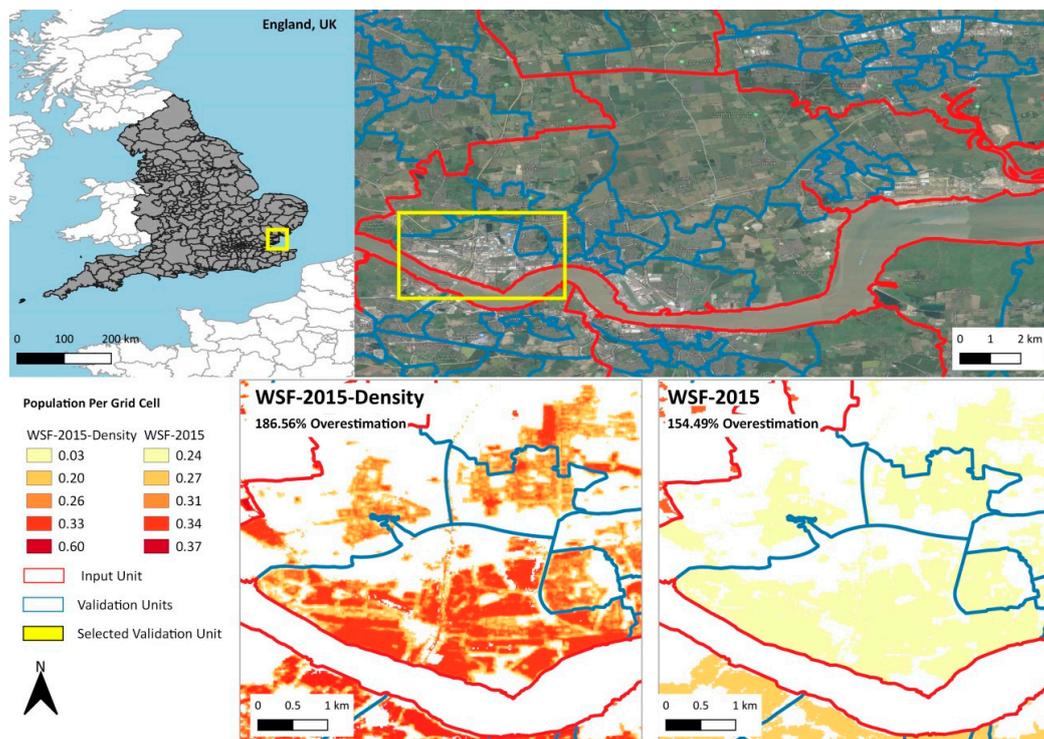


Figure 11. Influence of the building use in the population distribution results. Industrial areas capture large population counts resulting in large errors of overestimation within the validation units.

Similar accuracy limitations have been reported in the production of the GHS-POP and the HRSL population datasets [10,17]. Even when several local studies have demonstrated that information on the building use has the potential to improve population distribution results [61,62], this remains a major source of limitation in the production of global population datasets, as it is not possible to derive detailed semantic information on the building use through remote sensing methodologies. Population datasets such as LandScan and WorldPop integrate land use and land cover covariates to improve their results; however, as mentioned above, this introduces global transferability limitations and applicability restrictions.

For this reason, in this study, we began to analyse the relationship between the inherent characteristics of the underlying built-up environment and the performance of each covariate layer, as an alternative approach that could be used to minimise the errors introduced by the quality and lack of functional characterisation of the input covariate layers. Here, we introduced the SSC index as a globally transferable metric to categorise the input units in terms of the size and coverage of the underlying settlement objects. Our results clearly indicate that WSF-2015-Density layer distributes population with higher accuracies in regions with high SSC index values, reaching improvements up to ~30% over the WSF-2015 layer (Table 8). For regions with low and medium SSC index values, the performance of each covariate layer varies from country to country. Figure 12 shows the distribution of the SSC index values and the mean SSC index value for the “low” and “medium” SSC classes for each country.

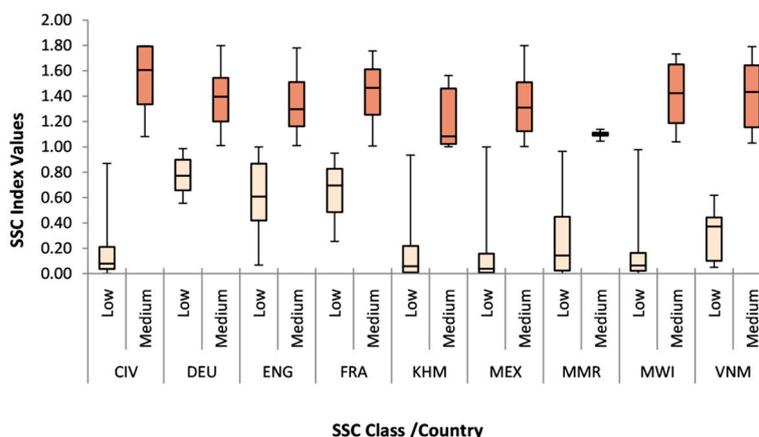


Figure 12. Boxplots of the distribution of the SSC index values for the “low” (yellow boxplots) and “medium” (green boxplots) classes for each country.

Focusing on the distribution of the “low” SSC class, countries where the WSF-2015 reported in average less RMSE values are also the countries where more than half of the input units reported SSC index values lower than 0.40. In other words, the SSC index values fell below the mean of the “low” SSC class that ranges from >0 to 1. For the “medium” SSC class, the distribution of the SSC index values among countries is relatively similar. The mixture of medium to highly populated cities and rural areas within these input units represent challenging modelling regions where further analyses are required to identify the particular circumstances where one layer outperforms the other.

Nevertheless, it is important to notice that the WSF-2015-Density layer performed better in all three classes of the SSC index for countries such as Germany, France and England, hence suggesting that the overall performance is largely driven by an accurate identification of building structures within the rural areas of each country. In this context, it is expected that limitations derived from the current underestimation of smaller settlements and isolated buildings can be overcome by the future integration of Sentinel-2 data in the production of future WSF datasets, due to its increased spatial resolution [40].

As a final note, it is important to mention that, even when the population distribution maps presented in this research have been produced using the most frequently employed population census data, the difficulties in the acquisition of the finest census data, the challenges in integrating census data with spatial boundaries and the uncertainties of population estimates based on statistical projections, are additional sources of errors and uncertainty limiting the accuracy of the population distribution models. Therefore, as stated by Doxsey-Whitfield et al. [6], acquiring up-to-date global population census data at the highest spatial detail possible should remain a priority for improving global population mapping.

5. Conclusions

The presented study focused on the cross-comparison of population distribution maps produced using the WSF-2015 and the experimental WSF-2015-Density layers. The main objective was to investigate if higher accuracies in population distribution mapping can be achieved using additional information on the build-up environment, such as the percentage of impervious surface, in comparison to the already established binary approach employed by other population datasets and their baseline settlement layers.

The results of the quantitative assessment showed that the overall accuracies between both covariate layers are comparably similar, with the best accuracy results reported for population distribution maps produced using the finest input census data. Our results indicate that, while both layers distribute the largest percentage of each country’s total population with estimation errors ranging from -25% to 25% , remaining limitations derived from: (i) the incomplete identification of

settlement pixels; and (ii) the lack of information on the building use, still introduce large errors of underestimation and errors overestimation in a considerable percentage of the population.

Notwithstanding these limitations, from a comparative point of view, our results have shown that population distribution maps produced on basis of the WSF-2015-Density layer provide a more realistic representation of the spatial distribution of the population, as the heterogeneous allocation of population counts prevents the appearance of artificial patterns between neighbouring administrative census units. Furthermore, it has been demonstrated that the WSF-2015-Density layer produces higher accuracies in high-density built-up environments and is capable to improve the estimation accuracies of the WSF-2015 layer up to ~30%, especially in those countries where a good percentage of building structures have been identified within the rural areas. The fact that the WSF-2015-Density layer is derived from remote sensing approaches that do not require a priori knowledge of the land cover makes it a strong suitable proxy capable to improve global population distribution methodologies, and, as it is not based on local relationships, it has no applicability restrictions in comparison to other existing products. Moreover, it provides global coverage and can be straightforward updated allowing time agreement with census population data, enabling the production of a consistent global population distribution dataset with higher accuracy and spatial resolution than those currently available.

One of the strengths of our study is the implementation of the SSC index, used to investigate the correlation between the built-up environment and the performance of each covariate layer. Our results suggest that higher accuracies in population disaggregation could be achieved with the correct preselection of the input covariate at the input unit level; however, to implement this preselection, additional research is still necessary, as the SSC index cannot provide a complete distinction between the covariate layers in areas with middle SSC index values.

However, in the light of these highly promising results, future research will focus on the validation and open release of the WSF-2015-Density layer, expanding the accuracy assessment of population mapping to other regions of the world, with special focus on arid and semi-arid areas, and comparing the results against other existing global population distribution datasets. Within this outlook, deeper research on the SSC index will also be included, to develop a methodology that can help minimise the inherent distribution errors derived from the quality and functional characterisation of the input covariates, as well as in the production of a new global population dataset.

Author Contributions: Conceptualisation, D.P.-L., T.E., M.M., A.S., A.J.T. and P.R.; Formal analysis, D.P.-L.; Investigation, D.P.-L.; Methodology, D.P.-L., F.B., T.E., A.H., M.M. and J.Z.; Supervision, T.E., W.H., A.J.T. and P.R.; Writing—original draft, D.P.-L.; and Writing—review and editing, D.P.-L., F.B., T.E., W.H., A.H., M.M., A.S., J.Z., C.K., S.D., A.J.T. and P.R.

Funding: This research received no external funding.

Acknowledgments: The authors would like to thank Kytt MacManus from the Center of International Earth Science Information Network (CIESIN) at Columbia University for the collection and provision of population census data and Sergey Voinov for sharing his expertise on dasymetric modelling.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. United Nations. *World Population Prospects 2019: Ten Key Findings*; United Nations, Department of Economic and Social Affairs, Population Division: New York, NY, USA, 2019.
2. United Nations. *World Population Prospects 2019: Highlights.ST/ESA/SER.A/423*; United Nations, Department of Economic and Social Affairs, Population Division: New York, NY, USA, 2019.
3. Balk, D.L.; Deichmann, U.; Yetman, G.; Pozzi, F.; Hay, S.I.; Nelson, A. Determining Global Population Distribution: Methods, Applications and Data. *Adv. Parasitol.* **2006**, *62*, 119–156. [[PubMed](#)]
4. Nieves, J.J.; Stevens, F.R.; Gaughan, A.E.; Linard, C.; Sorichetta, A.; Hornby, G.; Patel, N.N.; Tatem, A.J. Examining the correlates and drivers of human population distributions across low- and middle-income countries. *J. R. Soc. Interface* **2017**, *14*, 20170401. [[CrossRef](#)] [[PubMed](#)]

5. Center for International Earth Science Information Network - CIESIN - Columbia University; International Food Policy Research Institute - IFPRI; The World Bank; Centro Internacional de Agricultura Tropical - CIAT. *Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Population Density Grid*; NASA Socioeconomic Data and Applications Center (SEDAC): Palisades, NY, USA, 2011.
6. Doxsey-Whitfield, E.; MacManus, K.; Adamo, S.B.; Pistoiesi, L.; Squires, J.; Borkovska, O.; Baptista, S.R. Taking Advantage of the Improved Availability of Census Data: A First Look at the Gridded Population of the World, Version 4. *Papers Appl. Geogr.* **2015**, *1*, 226–234. [[CrossRef](#)]
7. Center for International Earth Science Information Network - CIESIN - Columbia University. *Gridded Population of the World, Version 4 (GPWv4): Population Count Adjusted to Match 2015 Revision of UN WPP Country Totals, Revision 11*; NASA Socioeconomic Data and Applications Center (SEDAC): Palisades, NY, USA, 2018.
8. Dobson, J.E.; Bright, E.A.; Coleman, P.R.; Durfee, R.C.; Worley, B.A. LandScan: a global population database for estimating populations at risk. *Photogramm. Eng. Remote Sens.* **2000**, *66*, 849–857.
9. Bhaduri, B.; Bright, E.; Coleman, P.; Urban, M.L. LandScan USA: a high-resolution geospatial and temporal modeling approach for population distribution and dynamics. *GeoJournal* **2007**, *69*, 103–117. [[CrossRef](#)]
10. Freire, S.; Doxsey-Whitfield, E.; MacManus, K.; Mills, J.; Pesaresi, M. Development of new open and free multi-temporal global population grids at 250 m resolution. In Proceedings of the AGILE, Helsinki, Finland, 14–17 June 2016.
11. Freire, S.; Kemper, T.; Pesaresi, M.; Florczyk, A.; Syrris, V. Combining GHSL and GPW to improve global population mapping. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 26–31 July 2015; pp. 2541–2543.
12. Lloyd, C.T.; Sorichetta, A.; Tatem, A.J. High resolution global gridded data for use in population studies. *Sci. Data* **2017**, *4*, 170001. [[CrossRef](#)]
13. Stevens, F.R.; Gaughan, A.E.; Linard, C.; Tatem, A.J. Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. *PLoS ONE* **2015**, *10*, e0107042. [[CrossRef](#)]
14. Sorichetta, A.; Hornby, G.M.; Stevens, F.R.; Gaughan, A.E.; Linard, C.; Tatem, A.J. High-resolution gridded population datasets for Latin America and the Caribbean in 2010, 2015, and 2020. *Sci. Data* **2015**, *2*, 150045. [[CrossRef](#)]
15. Gaughan, A.E.; Stevens, F.R.; Huang, Z.; Nieves, J.J.; Sorichetta, A.; Lai, S.; Ye, X.; Linard, C.; Hornby, G.M.; Hay, S.I.; et al. Spatiotemporal patterns of population in mainland China, 1990 to 2010. *Sci. Data* **2016**, *3*, 160005. [[CrossRef](#)]
16. Lloyd, C.T.; Chamberlain, H.; Kerr, D.; Yetman, G.; Pistoiesi, L.; Stevens, F.R.; Gaughan, A.E.; Nieves, J.J.; Hornby, G.; MacManus, K. Global spatio-temporally harmonised datasets for producing high-resolution gridded population distribution datasets. *Big Earth Data* **2019**, 1–32. [[CrossRef](#)]
17. Tiecke, T.G.; Liu, X.; Zhang, A.; Gros, A.; Li, N.; Yetman, G.; Kilic, T.; Murray, S.; Blankespoor, B.; Prydz, E.B. Mapping the world population one building at a time. *arXiv* **2017**, arXiv:1712.05839.
18. Elvidge, C.D.; Sutton, P.C.; Ghosh, T.; Tuttle, B.T.; Baugh, K.E.; Bhaduri, B.; Bright, E. A global poverty map derived from satellite data. *Comput. Geosci.* **2009**, *35*, 1652–1660. [[CrossRef](#)]
19. Noor, A.M.; Alegana, V.A.; Gething, P.W.; Tatem, A.J.; Snow, R.W. Using remotely sensed night-time light as a proxy for poverty in Africa. *Popul. Health Metrics* **2008**, *6*, 5. [[CrossRef](#)] [[PubMed](#)]
20. Barbier, E.B.; Hochard, J.P. Land degradation and poverty. *Nature Sustain.* **2018**, *1*, 623–631. [[CrossRef](#)]
21. Amoah, B.; Giorgi, E.; Heyes, D.J.; van Burren, S.; Diggle, P.J. Geostatistical modelling of the association between malaria and child growth in Africa. *Int. J. Health Geogr.* **2018**, *17*, 7. [[CrossRef](#)] [[PubMed](#)]
22. Dhewantara, P.W.; Mamun, A.A.; Zhang, W.-Y.; Yin, W.-W.; Ding, F.; Guo, D.; Hu, W.; Magalhães, R.J.S. Geographical and temporal distribution of the residual clusters of human leptospirosis in China, 2005–2016. *Sci. Rep.* **2018**, *8*, 16650. [[CrossRef](#)] [[PubMed](#)]
23. España, G.; Grefenstette, J.; Perkins, A.; Torres, C.; Campo Carey, A.; Diaz, H.; de la Hoz, F.; Burke, D.S.; van Panhuis, W.G. Exploring scenarios of chikungunya mitigation with a data-driven agent-based model of the 2014–2016 outbreak in Colombia. *Sci. Rep.* **2018**, *8*, 12201. [[CrossRef](#)]
24. Sorichetta, A.; Bird, T.J.; Ruktanonchai, N.W.; zu Erbach-Schoenberg, E.; Pezzulo, C.; Tejedor, N.; Waldock, I.C.; Sadler, J.D.; Garcia, A.J.; Sedda, L.; et al. Mapping internal connectivity through human migration in malaria endemic countries. *Sci. Data* **2016**, *3*, 160066. [[CrossRef](#)]

25. Linard, C.; Gilbert, M.; Snow, R.W.; Noor, A.M.; Tatem, A.J. Population distribution, settlement patterns and accessibility across Africa in 2010. *PLoS ONE* **2012**, *7*, e31743. [[CrossRef](#)]
26. Ajisegiri, B.; Andres, L.A.; Bhatt, S.; Dasgupta, B.; Echenique, J.A.; Gething, P.W.; Zabludovsky, J.G.; Joseph, G. Geo-spatial modeling of access to water and sanitation in Nigeria. *J. Water Sanit. Hyg. Dev.* **2019**, *9*, 258–280. [[CrossRef](#)]
27. Linard, C.; Kabaria, C.W.; Gilbert, M.; Tatem, A.J.; Gaughan, A.E.; Stevens, F.R.; Sorichetta, A.; Noor, A.M.; Snow, R.W. Modelling changing population distributions: an example of the Kenyan Coast, 1979–2009. *Int. J. Digital Earth* **2017**, *10*, 1017–1029. [[CrossRef](#)] [[PubMed](#)]
28. Weber, E.M.; Seaman, V.Y.; Stewart, R.N.; Bird, T.J.; Tatem, A.J.; McKee, J.J.; Bhaduri, B.L.; Moehl, J.J.; Reith, A. Census-independent population mapping in northern Nigeria. *Remote Sens. Environ.* **2018**, *204*, 786–798. [[CrossRef](#)] [[PubMed](#)]
29. Brown, S.; Nicholls, R.J.; Goodwin, P.; Haigh, I.; Lincke, D.; Vafeidis, A.; Hinkel, J. Quantifying land and people exposed to sea-level rise with no mitigation and 1.5 C and 2.0 C rise in global temperatures to year 2300. *Earth's Future* **2018**, *6*, 583–600. [[CrossRef](#)]
30. Aubrecht, C.; Özceylan, D.; Steinnocher, K.; Freire, S. Multi-level geospatial modeling of human exposure patterns and vulnerability indicators. *Nat. Hazards* **2013**, *68*, 147–163. [[CrossRef](#)]
31. Maas, P.; Iyer, S.; Gros, A.; Park, W.; McGorman, L.; Nayak, C.; Dow, P.A. Facebook Disaster Maps: Aggregate Insights for Crisis Response and Recovery. In Proceedings of the 16th International Conference on Information Systems for Crisis Response and Management (ISCRAM), Valencia, Spain, 19–22 May 2019; p. 3173.
32. Serrano Gine, D.; Russo, A.; Brandajs, F.; Perez Albert, M. Characterizing European urban settlements from population data: A cartographic approach. *Cartogr. Geogr. Inform. Sci.* **2016**, *43*, 442–453. [[CrossRef](#)]
33. Su, M.D.; Lin, M.C.; Hsieh, H.I.; Tsai, B.W.; Lin, C.H. Multi-layer multi-class dasymetric mapping to estimate population distribution. *Sci. Total Environ.* **2010**, *408*, 4807–4816. [[CrossRef](#)]
34. Reed, F.; Gaughan, A.; Stevens, F.; Yetman, G.; Sorichetta, A.; Tatem, A. Gridded population maps informed by different built settlement products. *Data* **2018**, *3*, 33. [[CrossRef](#)]
35. Hay, S.I.; Noor, A.M.; Nelson, A.; Tatem, A.J. The accuracy of human population maps for public health application. *Trop Med. Int. Health* **2005**, *10*, 1073–1086. [[CrossRef](#)]
36. Nagle, N.N.; Buttenfield, B.P.; Leyk, S.; Spielman, S. Dasymetric modeling and uncertainty. *Ann. Assoc. Am. Geogr.* **2014**, *104*, 80–95. [[CrossRef](#)]
37. Pesaresi, M.; Ehrlich, D.; Ferri, S.; Florczyk, A.; Freire, S.; Halkia, M.; Julea, A.; Kemper, T.; Soille, P.; Syrris, V. *Operating procedure for the production of the Global Human Settlement Layer from Landsat data of the epochs 1975, 1990, 2000, and 2014*; Joint Research Centre: Ispra, Italy, 2016; pp. 1–62. ISBN 978-92-79-55012-6.
38. Esch, T.; Heldens, W.; Hirner, A.; Keil, M.; Marconcini, M.; Roth, A.; Zeidler, J.; Dech, S.; Strano, E.; Sensing, R. Breaking new ground in mapping human settlements from space—The Global Urban Footprint. *ISPRS J. Photogramm.* **2017**, *134*, 30–42. [[CrossRef](#)]
39. Esch, T.; Marconcini, M.; Felbier, A.; Roth, A.; Heldens, W.; Huber, M.; Schwinger, M.; Taubenböck, H.; Müller, A.; Dech, S. Urban footprint processor—Fully automated processing chain generating settlement masks from global data of the TanDEM-X mission. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 1617–1621. [[CrossRef](#)]
40. Esch, T.; Bachofer, F.; Heldens, W.; Hirner, A.; Marconcini, M.; Palacios-Lopez, D.; Roth, A.; Üreyen, S.; Zeidler, J.; Dech, S. Where we live—A summary of the achievements and planned evolution of the global urban footprint. *Remote Sens.* **2018**, *10*, 895. [[CrossRef](#)]
41. Steinnocher, K.; De Bono, A.; Chatenoux, B.; Tiede, D.; Wendt, L. Estimating urban population patterns from stereo-satellite imagery. *Eur. J. Remote Sens.* **2019**, *52*, 1–14. [[CrossRef](#)]
42. Merkens, J.-L.; Vafeidis, A. Using Information on Settlement Patterns to Improve the Spatial Distribution of Population in Coastal Impact Assessments. *Sustainability* **2018**, *10*, 3170. [[CrossRef](#)]
43. Marconcini, M.; Metz-Marconcini, A.; Üreyen, S.; Palacios-Lopez, D.; Hanke, W.; Bachofer, F.; Zeidler, J.; Esch, T.; Gorelick, N.; Kakarla, A.; et al. Outlining where humans live - The World Settlement Footprint 2015. *Sci. Data* **2019**. [Submitted].
44. Bauer, M.E.; Löffelholz, B.C.; Wilson, B. *Estimating and mapping impervious surface area by regression analysis of Landsat imagery*; CRC Press: Boca Raton, FL, USA, 2007; pp. 31–48.

45. Azar, D.; Graesser, J.; Engstrom, R.; Comenetz, J.; Leddy, R.M.; Schechtman, N.G.; Andrews, T. Spatial refinement of census population distribution using remotely sensed estimates of impervious surfaces in Haiti. *Int. J. Remote Sens.* **2010**, *31*, 5635–5655. [CrossRef]
46. Lu, D.; Weng, Q.; Li, G. Residential population estimation using a remote sensing derived impervious surface approach. *Int. J. Remote Sens.* **2006**, *27*, 3553–3570. [CrossRef]
47. Li, G.; Weng, Q. Using Landsat ETM+ imagery to measure population density in Indianapolis, Indiana, USA. *Photogramm. Eng.* **2005**, *71*, 947–958. [CrossRef]
48. Marconcini, M.; Metz, A.; Zeidler, J.; Esch, T. Urban monitoring in support of sustainable cities. In *2015 Joint Urban Remote Sensing Event (JURSE)*; IEEE: Piscataway, NJ, USA, 2015; pp. 1–4.
49. Esch, T.; Üreyen, S.; Zeidler, J.; Metz–Marconcini, A.; Hirner, A.; Asamer, H.; Tum, M.; Böttcher, M.; Kuchar, S.; Svaton, V.; et al. Exploiting big earth data from space – first experiences with the timescan processing chain. *Big Earth Data* **2018**, *2*, 36–55. [CrossRef]
50. Ritchie, H.; Roser, M. Urbanization. Available online: <https://ourworldindata.org/urbanization> (accessed on 13 May 2019).
51. Taw, N.P. *The 2014 Myanmar Population and Housing Census. Highlights of the Main Results*; M.o.I.a.P. Department of Population: Myanmar, 29 May 2015; pp. 1–47.
52. GeoNode. Available online: <http://geonode.themimu.info/> (accessed on 10 February 2019).
53. Li, L.; Lu, D. Mapping population density distribution at multiple scales in Zhejiang Province using Landsat Thematic Mapper and census data. *Int. J. Remote Sens.* **2016**, *37*, 4243–4260. [CrossRef]
54. Bai, Z.; Wang, J.; Wang, M.; Gao, M.; Sun, J. Accuracy Assessment of Multi-Source Gridded Population Distribution Datasets in China. *Sustainability* **2018**, *10*, 1363. [CrossRef]
55. Tatem, A.J.; Noor, A.M.; von Hagen, C.; Di Gregorio, A.; Hay, S.I. High resolution population maps for low income nations: Combining land cover and census in East Africa. *PLoS ONE* **2007**, *2*, e1298. [CrossRef] [PubMed]
56. Anderson-Sprecher, R. Model comparisons and R2. *Am. Stat.* **1994**, *48*, 113–117.
57. Mennis, J.; Hultgren, T. Intelligent dasymetric mapping and its application to areal interpolation. *Cartogr. Geogr. Inform. Sci.* **2006**, *33*, 179–194. [CrossRef]
58. Esch, T.; Marconcini, M.; Marmanis, D.; Zeidler, J.; Elsayed, S.; Metz, A.; Müller, A.; Dech, S. Dimensioning urbanization – An advanced procedure for characterizing human settlement properties and patterns using spatial network analysis. *Appl. Geogr.* **2014**, *55*, 212–228. [CrossRef]
59. Qi, Y.; Wu, J. Effects of changing spatial resolution on the results of landscape pattern analysis using spatial autocorrelation indices. *Landscape Ecol.* **1996**, *11*, 39–49. [CrossRef]
60. Goodwin, L.D.; Leech, N.L. Understanding correlation: Factors that affect the size of r. *J. Exper. Educ.* **2006**, *74*, 249–266. [CrossRef]
61. Biljecki, F.; Ohori, K.A.; Ledoux, H.; Peters, R.; Stoter, J. Population estimation using a 3D city model: A multi-scale country-wide study in the Netherlands. *PLoS ONE* **2016**, *11*, e0156808. [CrossRef]
62. Goerlich, F. A volumetric approach to spatial population disaggregation using a raster build-up layer, land use/land cover databases (SIOSE) and LIDAR remote sensing data. *Revista de Teledetección* **2016**, 147–163.

