

Appendix A

Supporting Information, for:

Assessing the accuracy of census-independent small area modelled population datasets

Extended Results

Supporting Datasets

The analyses described in the main text concern comparisons between the Zambia bottom-up modelled population count estimates (Dooley et al., 2021; WorldPop, 2020) and enumerated population counts from the Zambia 2022 census (Zambia Statistics Agency, 2024). In these comparisons, the modelled population estimates are initially compared directly against the enumerated census counts. The subsequent analyses focus on two grid cell subsets of the modelled estimates, which we term “conservative” and “expanded”. These subsets are identified based on the location of residential census households. In all these analyses, the modelled estimates are compared against the census population counts published for each ward, district and province nationally. The modelled population estimates and census population counts are provided as supporting datasets in an Excel spreadsheet (**Dataset S1**), as described at the end of the SI.

Supporting Figures and Tables

The additional Figures and Tables provided in the SI are intended to support the results described in the main text. These include:

- Comparisons between modelled population estimates with census counts at ward-level, stratified by province (**Figure A.1**)
- Model evaluation statistics (R (Pearson), RMSE, MAE, MAPE, MedAE, bias, imprecision) with stratification by province (**Table A.1**)
- Summaries of the proportion of administrative units for which the census counts were within the 95% CIs of the modelled estimates (**Table A.2** – districts, **Table A.3** – wards), with stratification by province

The conservative and expanded grid cell subsets are defined based on the location of census residential household locations. Additional Figures are provided in the SI to support the definition of these subsets:

- **Figure A.2** shows the (1) number of grid cells and (2) modelled estimates of population, for (i) all grid cells, (ii) the expanded subset, and (iii) the conservative subset.
- **Figure A.4** shows how the location of grid cells with modelled population estimates (lower axis) maps to locations of residential census households (upper axis) nationally, with summaries per province provided in **Figure A.5**.

A limitation of the analyses to compare modelled estimates (all grid cells and grid cell subsets) and census counts, concerns the presence of grid cells with residential households at the time of the census, but without modelled population estimates due to an absence of building footprints. As discussed in the main text, the gridded population estimates nationally include 2,273,498 grid cells with estimated population, corresponding to a national population estimate of 21,657,131 (95% CI: 20,912,731 - 22,474,310). In comparison, the 2022 census population was 19,693,423 (Zambia Statistics Agency, 2024), and there were 1,525,866 grid cells with residential households recorded in the census (WorldPop and ZamStats, 2024).

As shown in **Figure A.4**, of the 1,525,866 grid cells nationally with residential census households, 1,145,112 of these also have population predicted in the modelled estimates, which we termed the “conservative” subset (green flow). The remaining 380,754 grid cells with residential census households do not have modelled population estimates (purple flow). When we relax the definition of grid cells with residential census households to also consider the eight surrounding grid cells (identified through the application of a 3x3 spatial filter), an additional 4,807,710 grid cells are identified as potentially having residential households. Of these 4,807,710 grid cells, approximately 12.5% (n=602,584) are grid cells with modelled population estimates (yellow flow), whilst the remaining grid cells do not have building footprints (red flow, n=4,205,126). The green and yellow flows together represent the “expanded” subset (1,747,697 grid cells). The blue flow represents the 525,721 grid cells that have modelled population estimates but are not included in either the conservative or expanded subset, due to an absence of residential census households. Grid cells in the blue flow are likely to: (1) have only non-residential or uninhabited buildings, (2) have building footprints that are false positives, or (3) have residential buildings but these are missing from the census residential household definition. **Figure A.5** builds on **Figure A.4** to show how the classification of grid cells in the modelled population estimates dataset and the census residential households dataset compare on a province-by-province basis. Relatively consistent patterns are seen between provinces.

The situation of grid cells with residential households in the census, but without modelled population estimates, can largely be attributed to the temporal differences between the 2022 census and the year of the satellite imagery from which the building footprints were extracted. Nationally, 92% of grid cells with building footprints were extracted from satellite imagery dating from 2017 (22.9%), 2018 (42.5%) or 2019 (26.8%), with the remaining building footprints extracted from satellite imagery from preceding years, with the earliest imagery dating from 2010. Thus, there is a minimum of 3 years elapsed between the most recent satellite imagery used in the building footprint extraction, and the census. **Figure A.3** shows the year of the satellite imagery from which the building footprints used in the population modelling were extracted, summarized in terms of the count of grid cells for each year (2010-2019), with stratification by province.

As described in the main text, we conducted an additional analysis to identify those grid cells with residential households at the time of the census, but for which no modelled population estimates were predicted (n=380,754 nationally). **Table A.8** provides a summary of the count of grid cells with modelled population estimates (building footprints), residential census households, and residential census households but no modelled population estimates, summarised both nationally and per province. We estimated the portion of the census population count that is potentially resident in those grid cells without building footprints. For each ward, the count of census households in the subset of grid cells was multiplied by the mean household size (estimated based on the census count of population divided by the count of census households), to estimate the census population associated with those grid cells. Across the 380,754 grid cells nationally with residential census households but no modelled population estimates, we estimated the census population resident in those grid cells to be 2,311,349, which when subtracted from the national census population total, gives an adjusted national census population of 17,382,074 (**Table A.8**).

Additional analyses comparing the subsets of modelled estimates with the adjusted census comparator are described in the main text. The following SI Figures and Tables support this additional analysis:

- **Figure A.6** shows the modelled population estimates for expanded and conservative subsets, compared against adjusted census counts for province, district and wards.
- **Figure A.7** provides scatterplots for modelled ward estimates based on the conservative subset, plotted against adjusted census counts, with stratification by province.
- **Table A.4** provides model evaluation statistics (R (Pearson), RMSE, MAE, MAPE, MedAE, bias, imprecision) comparing the modelled population estimates based on the expanded

- and conservative subsets, and the adjusted census totals, for all districts and wards nationally. **Table A.5** provides the same model evaluation statistics with stratification by province.
- Summaries of the proportion of administrative units for which the adjusted census totals were within the 95% CIs of the modelled estimates subsets are provided for districts and wards in **Table A.6**. Additional summaries for the conservative subset of estimates are provided in **Table A.7**, which quantify the percentage of spatial units in which the lower 95% CI exceeds the adjusted census count, and the upper 95% CI is less than the adjusted census count. Summaries are provided for districts and wards nationally, as well as with stratification by province.

Figures

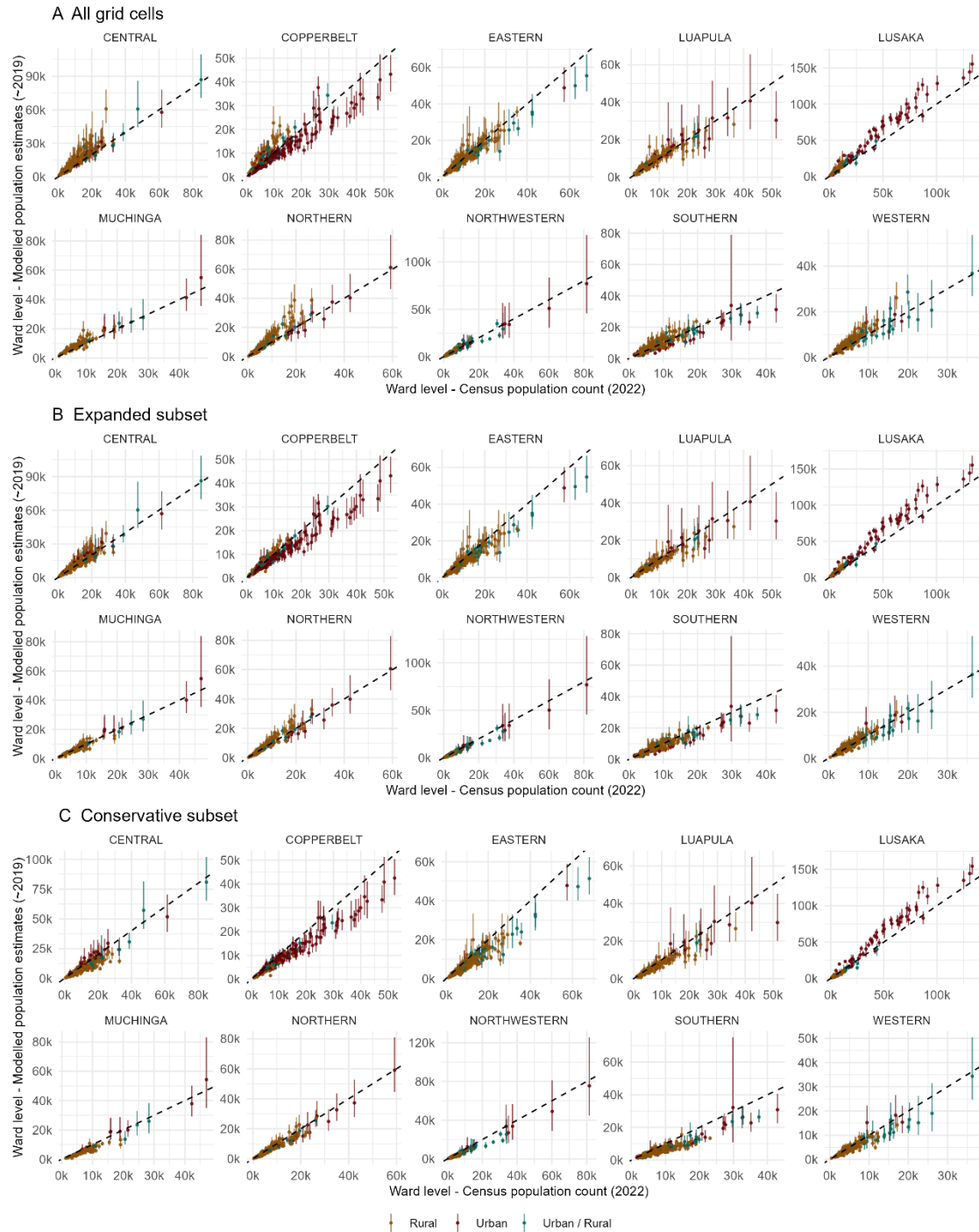


Fig. A.1. Modelled population estimates with 95% credible intervals (CIs) plotted against census grid counts for each ward, stratified by province. Parts A-C show modelled estimates based on (A) all grid cells, (B) the expanded subset, and (C) the conservative subset. The diagonal black line ($y=x$) shows a one-to-one relationship between the modelled and census-enumerated counts. The axis limits

vary between plots to accommodate the wide range of values in ward population. The rural/urban designation of each ward, as used in the census, is also shown.

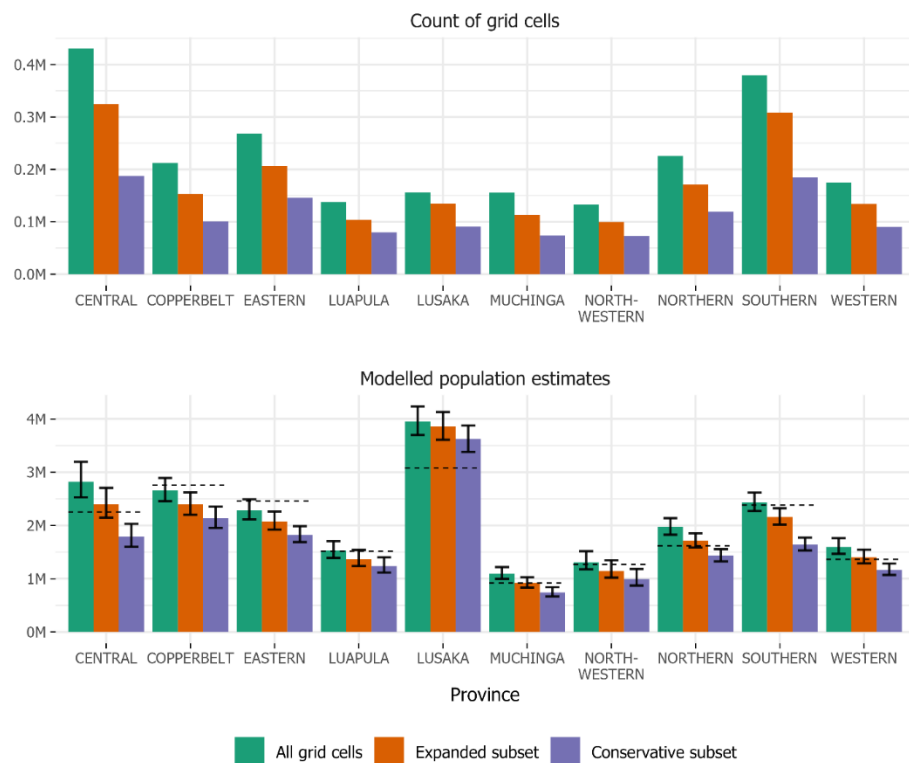


Fig. A.2. For each province, the count of grid cells with modelled population estimates (upper) and the modelled estimates of population (lower) are shown, for all grid cells (green), the expanded subset of grid cells (orange) and the conservative subset of grid cells (purple). In the lower plot, the green/orange/purple bars show the mean value of the modelled population estimates, with the 95% CIs shown by the error bars, and the 2022 enumerated census count shown by the dotted line.

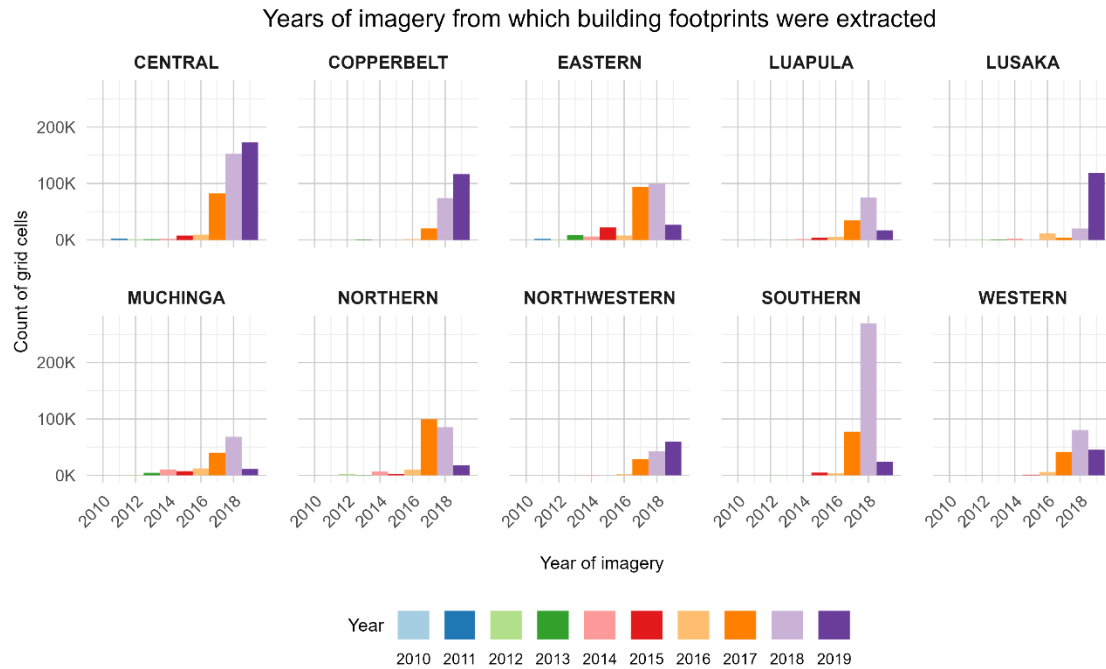


Fig. A.3. The year of the satellite imagery, from which Ecopia building footprints were extracted, quantified in terms of the count of grid cells with building footprints, stratified by province.

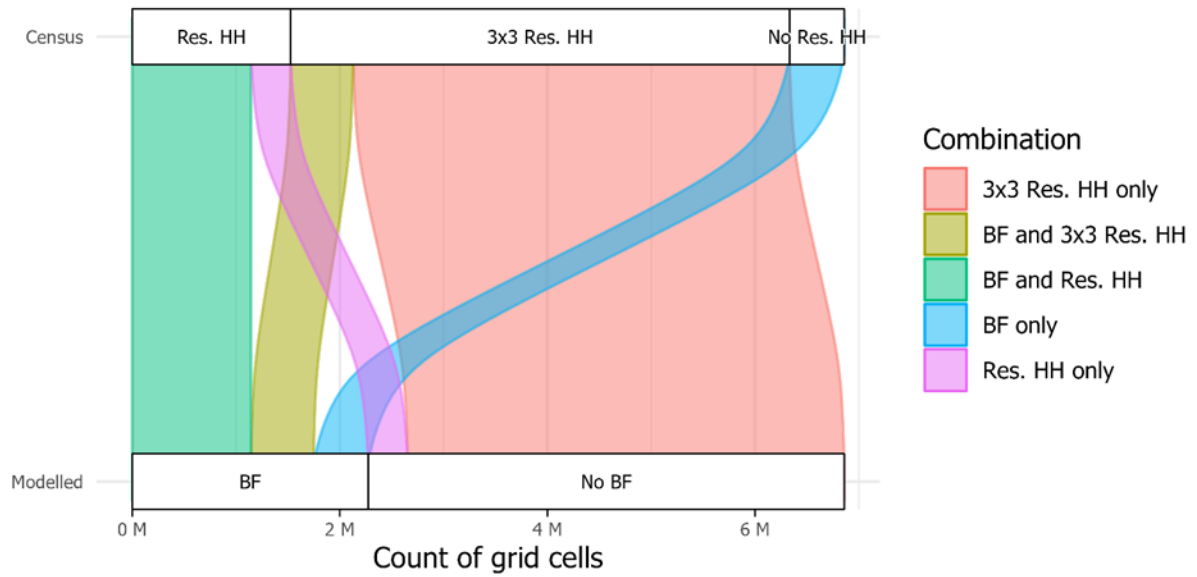


Fig. A.4. The count of grid cells with building footprints (BFs), and thus modelled population estimates (lower axis), and the count of grid cells with residential census households (HHs; upper axis). The spatial agreement between these classifications are shown by the flows between the upper and lower axes. The conservative subset of grid cells is represented by the green flow, the expanded subset is represented by the combination of the green and yellow flows. The purple flow represents grid cells with residential census households but with no building footprints and therefore no modelled population estimates. The blue flow represents grid cells with building footprints but no residential census households.

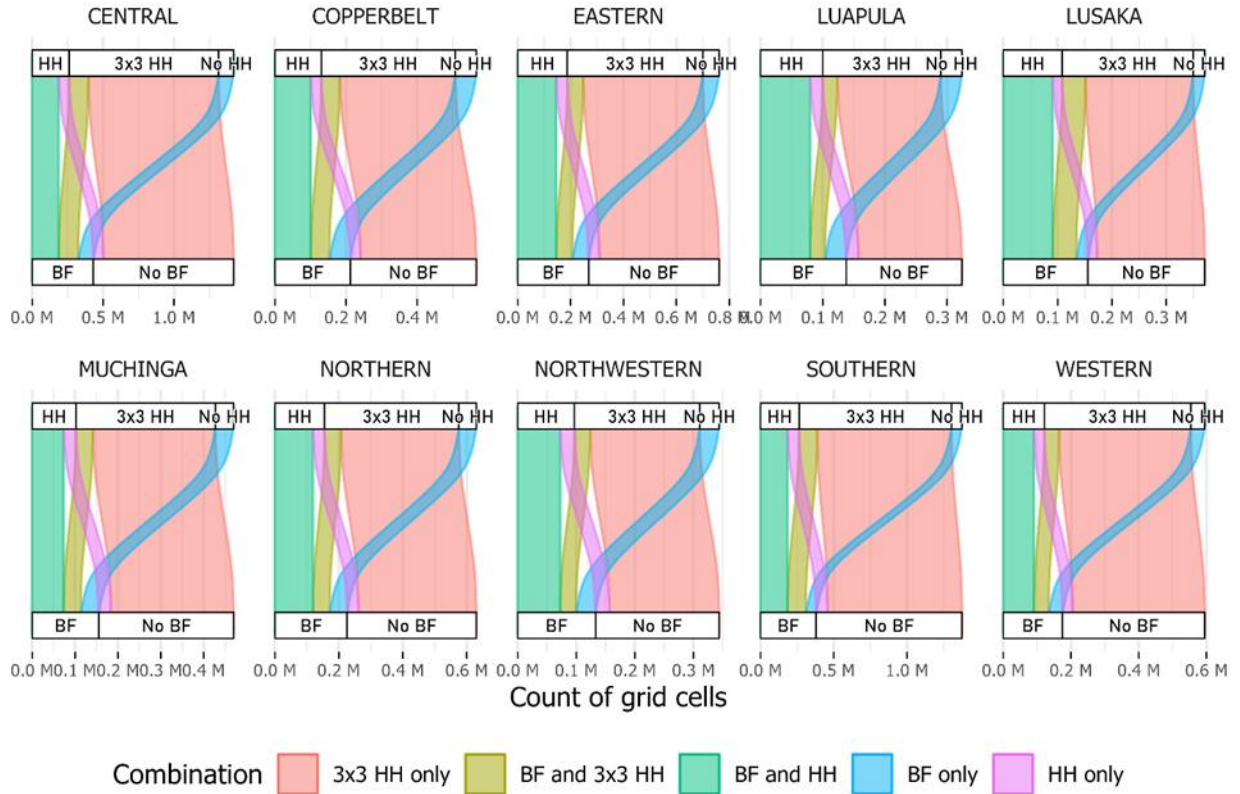


Fig. A.5. For each province, the count of grid cells with building footprints (BFs), and thus modelled population estimates (lower axis), and the count of grid cells with residential census households (HHs; upper axis). The spatial agreement between these classifications are shown by the flows between the upper and lower axes. The conservative subset of grid cells is represented by the green flow, the expanded subset is represented by the combination of the green and yellow flows. The purple flow represents grid cells with residential census households but with no building footprints and therefore no modelled population estimates. The blue flow represents grid cells with building footprints but no residential census households.

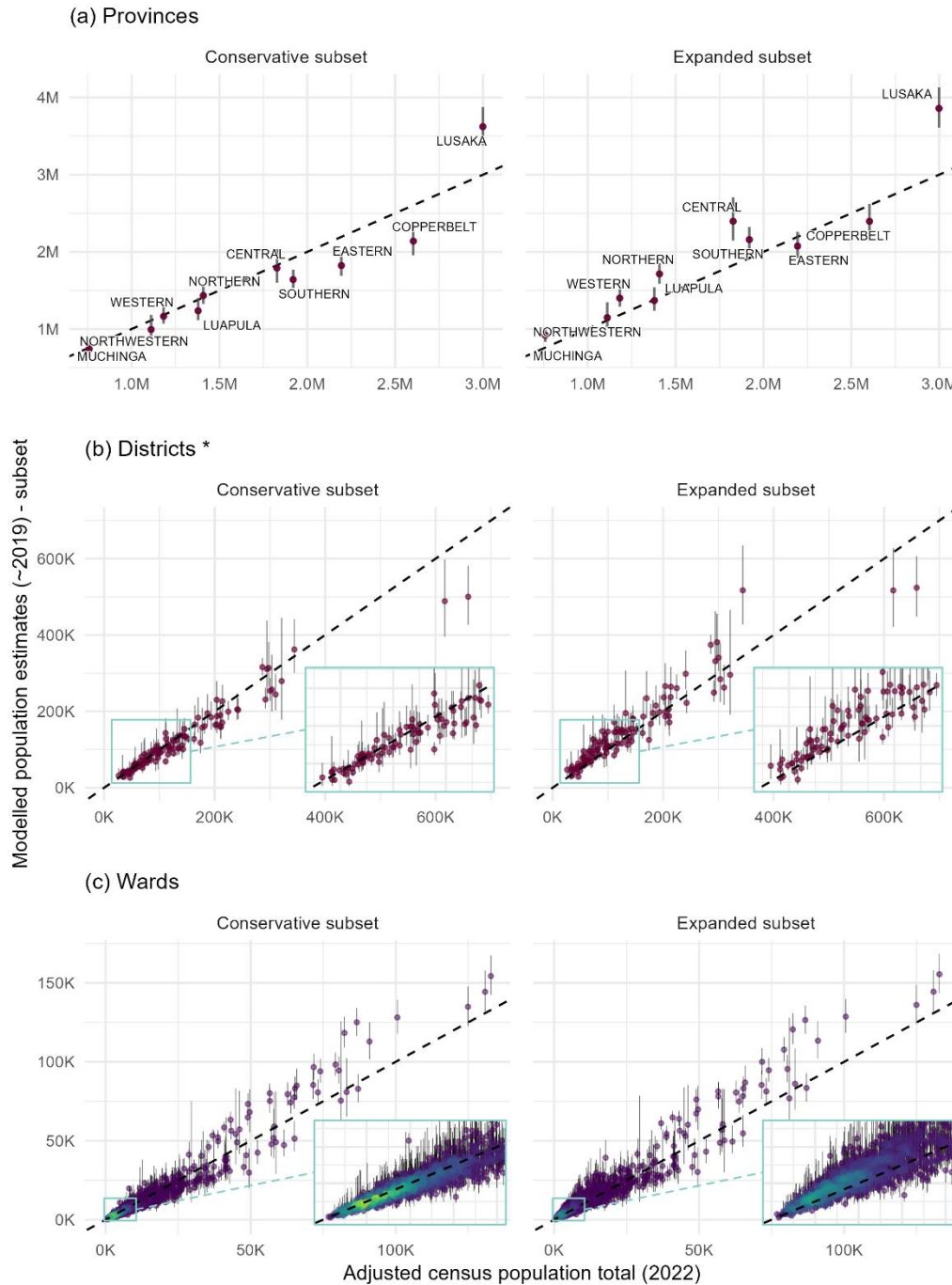


Fig. A.6. Modelled population estimates with 95% credible intervals (CIs) for expanded and conservative subsets, plotted against adjusted census counts for each (a) province ($n=10$), (b) district ($n=116$) and (c) ward ($n=1,858$) nationally. The district of Lusaka is omitted from the district plots in (b) due to being the only district with population exceeding 1 million (see Figure 2 of main text). The diagonal dashed line ($y=x$) on each plot shows a one-to-one relationship between the modelled estimates and adjusted census totals. The points on the ward plot are coloured based on

the density of points, with brighter colours (yellow/green) showing the greatest density of points. See Fig. A.7 for ward plots stratified by province.

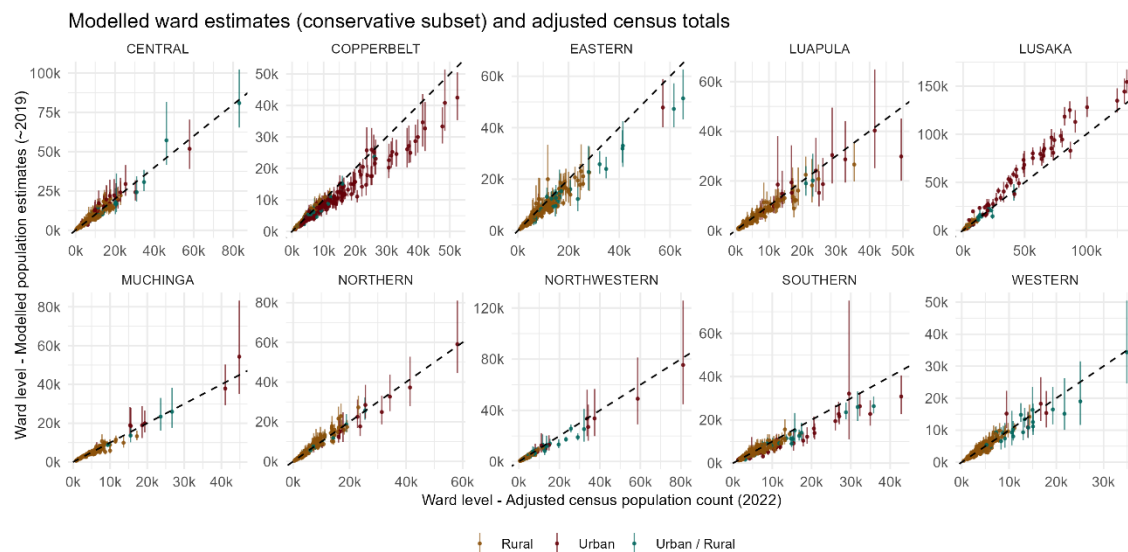


Fig. A.7. Modelled population estimates (conservative subset) with 95% credible intervals (CIs) plotted against adjusted census counts for each ward, stratified by province. The diagonal black line ($y=x$) shows a one-to-one relationship between the modelled and census-enumerated counts. The axis limits vary between plots to accommodate the wide range of values in ward population. The rural/urban designation of each ward, as used in the census, is also shown.

Tables

Table A.1. Model evaluation statistics, comparing modelled estimates against census enumeration at ward-level, with stratification by province.

Metric	Ward population									
	Central	Copper-belt	Eastern	Luapula	Lusaka	Muchinga	Northern	North-western	Southern	Western
n wards	180	253	206	169	118	104	180	167	254	227
All grid cells										
R (Pearson)	0.925	0.919	0.961	0.935	0.986	0.968	0.935	0.981	0.918	0.919
RMSE	5,383.4	3,753.8	2,986.1	2,865.4	11,548.5	2,492.1	3,774.1	1,932.6	2,582.5	2,101.1
MAE	3,508.0	2,676.4	1,943.8	1,626.2	7,633.1	1,874.6	2,443.4	1,185.6	1,796.3	1,485.2
MAPE [%]	34.4	28.3	17.4	20.6	37.5	26.6	28.7	18.2	21.1	26.2
MedAE	2,028.3	1,743.7	1,259.0	1,003.1	4,271.0	1,376.2	1,282.0	689.7	1,111.7	1,084.5
Bias	3,122.2	-434.4	-868.7	49.0	7,293.4	1,660.3	1,951.5	192.8	180.1	968.2
Imprecision	4,397.7	3,735.9	2,863.9	2,873.5	8,992.1	1,867.5	3,239.4	1,928.8	2,581.3	1,868.9
Expanded subset of grid cells										
R (Pearson)	0.954	0.966	0.965	0.944	0.987	0.973	0.968	0.987	0.944	0.947
RMSE	3,270.2	3,221.7	3,422.3	2,890.6	11,212.7	1,696.1	2,087.6	1,926.0	2,436.8	1,435.6
MAE	2,184.3	2,073.9	2,348.3	1,584.1	6,899.7	1,118.5	1,391.9	1,094.7	1,681.6	1,008.3
MAPE [%]	18.6	17.6	20.0	16.8	27.9	13.2	16.0	15.4	17.6	17.8
MedAE	1,314.7	1,206.3	1,614.6	903.7	2,359.2	648.3	899.5	622.3	1,055.4	670.3
Bias	749.2	-1,473.3	-1,883.3	-897.2	6,487.4	-34.3	501.5	-783.8	-904.0	111.1
Imprecision	3,192.1	2,870.7	2,864.5	2,756.0	9,184.3	1,703.9	2,032.1	1,764.6	2,267.3	1,434.5
Conservative subset of grid cells										
R (Pearson)	0.941	0.982	0.956	0.948	0.987	0.966	0.976	0.986	0.944	0.949
RMSE	4,295.6	3,569.1	4,525.0	3,228.5	10,148.7	2,531.6	1,998.2	2,615.6	3,858.9	1,738.5
MAE	3,087.3	2,529.7	3,264.7	1,879.5	6,253.3	2,027.0	1,446.9	1,740.9	3,019.2	1,153.1
MAPE [%]	25.8	22.8	27.9	20.3	24.8	26.2	17.9	25.1	31.0	18.5
MedAE	2,126.9	1,748.5	2,413.2	1,026.5	2,341.0	1,689.8	1,083.0	1,216.9	2,354.9	741.1
Bias	-2,621.4	-2,486.6	-3,107.1	-1,675.2	4,480.4	-1,738.5	-1,056.8	-1,708.5	-2,941.8	-928.3
Imprecision	3,412.5	2,565.5	3,297.6	2,768.1	9,145.0	1,849.3	1,700.6	1,986.4	2,502.2	1,473.2

Table A.2. The percentage of districts for which the census enumerated population counts are within the 95% CIs of the modelled population estimates, calculated from (i) all grid cells, (ii) the expanded subset, and (iii) the conservative subset. For those districts in which the lower 95% CI exceeds the census count, and the upper 95% CI is less than the census count also shown. Values are shown for all spatial units nationally, as well as with stratification by province.

		All grid cells			Expanded subset			Conservative subset		
		within 95% CIs	L95% CI > census	U95% CI < census	within 95% CIs	L95% CI > census	U95% CI < census	within 95% CIs	L95% CI > census	U95% CI < census
National (n=116)		63.8	29.3	6.9	75.0	5.2	19.8	44.0	0.9	55.2
Province	Central (n=11)	72.7	27.3	0.0	90.9	9.1	0.0	72.7	0.0	27.3
	Copperbelt (n=10)	50.0	30.0	20.0	60.0	0.0	40.0	10.0	0.0	90.0
	Eastern (n=15)	66.7	6.7	26.7	60.0	0.0	40.0	33.3	0.0	66.7
	Luapula (n=12)	75.0	16.7	8.3	83.3	0.0	16.7	66.7	0.0	33.3
	Lusaka (n=6)	33.3	66.7	0.0	66.7	33.3	0.0	83.3	16.7	0.0
	Muchinga (n=8)	50.0	50.0	0.0	87.5	0.0	12.5	25.0	0.0	75.0
	Northern (n=12)	58.3	41.7	0.0	91.7	8.3	0.0	50.0	0.0	50.0
	Northwestern (n=11)	81.8	18.2	0.0	63.6	0.0	36.4	45.5	0.0	54.5
	Southern (n=15)	80.0	13.3	6.7	60.0	0.0	40.0	13.3	0.0	86.7
	Western (n=16)	50.0	50.0	0.0	87.5	12.5	0.0	62.5	0.0	37.5

Table A.3. The percentage of wards for which the census enumerated population counts are within the 95% CIs of the modelled population estimates, calculated from (i) all grid cells, (ii) the expanded subset, and (iii) the conservative subset. For those wards in which the census enumerated counts are outside the 95% CIs, the percentage of wards in which the lower 95% CI exceeds the census count, and the upper 95% CI is less than the census count also shown. Values are shown for all spatial units nationally, as well as with stratification by province.

		All grid cells			Expanded subset			Conservative subset		
		within 95% CIs	L95% CI > census	U95% CI < census	within 95% CIs	L95% CI > census	U95% CI < census	within 95% CIs	L95% CI > census	U95% CI < census
National (n=1,858)		53.9	35.4	10.7	65.8	12.5	21.7	44.4	2.5	53.1
Province	Central (n=180)	60.0	39.4	0.6	80.6	11.1	8.3	58.9	1.1	40.0
	Copperbelt (n=253)	44.3	28.1	27.7	58.1	7.5	34.4	35.6	0.0	64.4
	Eastern (n=206)	68.9	10.2	20.9	63.1	2.4	34.5	42.7	0.0	57.3
	Luapula (n=169)	65.1	27.2	7.7	71.6	7.1	21.3	62.7	0.6	36.7
	Lusaka (n=118)	33.1	63.6	3.4	51.7	43.2	5.1	54.2	28.8	16.9
	Muchinga (n=104)	43.3	54.8	1.9	74.0	11.5	14.4	26.0	0.0	74.0
	Northern (n=180)	47.2	45.0	7.8	70.0	17.8	12.2	57.2	0.0	42.8
	Northwestern (n=167)	67.1	26.3	6.6	62.3	5.4	32.3	33.5	0.0	66.5
	Southern (n=254)	59.4	28.0	12.6	66.5	7.5	26.0	24.4	0.0	75.6
	Western (n=227)	43.2	52.9	4.0	63.0	23.3	13.7	54.2	4.0	41.9

Table A.4. Model evaluation statistics, comparing the mean value of the modelled estimates from the expanded and conservative grid cell subsets, against adjusted census totals for districts and wards.

Statistic	District-level		Ward-level	
	Expanded subset	Conservative subset	Expanded subset	Conservative subset
R (Pearson)	0.984	0.985	0.960	0.967
RMSE	67,257.0	58,013.7	4,039.5	3,439.7
MAE	30,287.4	22,105.5	2,284.8	1,678.6
MAPE [%]	21.4	12.3	28.7	16.2
MedAE	19,360.1	8,803.4	1,327.1	722.8
Bias	17,679.2	-6,882.9	1,103.8	-429.7
Imprecision	65,173.3	57,853.9	3,886.8	3,413.7

Table A.5. Model evaluation statistics, comparing the mean value of the modelled estimates based on the conservative and expanded subsets, against adjusted census population totals at ward-level, with stratification by province.

Metric	Ward population									
	Central	Copper-belt	Eastern	Luapula	Lusaka	Muchinga	Northern	North-western	Southern	Western
Expanded subset of grid cells										
R (Pearson)	0.946	0.949	0.971	0.950	0.987	0.973	0.953	0.987	0.907	0.926
RMSE	4687.1	3400.9	2535.1	2451.5	11528.4	2294.1	3006.3	1675.8	2776.0	1893.6
MAE	3326.8	2388.3	1560.9	1348.8	7600.8	1778.1	2072.9	1047.2	2131.5	1416.0
MAPE [%]	43.2	27.7	16.3	18.6	41.3	33.4	29.2	19.6	32.6	29.9
MedAE	2513.8	1650.0	872.8	773.5	3851.5	1512.8	1291.0	662.1	1645.6	1081.7
Bias	3160.7	-821.6	-578.8	-63.9	7290.7	1538.0	1698.4	216.8	941.0	962.9
Imprecision	3470.8	3306.7	2474.2	2458.0	8968.3	1710.4	2487.5	1666.7	2616.8	1634.1
Conservative subset of grid cells										
R (Pearson)	0.973	0.980	0.973	0.958	0.988	0.979	0.977	0.990	0.961	0.958
RMSE	2209.7	3249.8	3201.8	2500.4	10232.8	1497.3	1560.4	1794.0	2327.6	1194.1
MAE	1503.4	2034.6	2134.5	1293.5	6010.0	826.6	1013.2	910.1	1482.7	744.3
MAPE [%]	15.3	17.1	19.7	14.5	23.1	10.7	13.4	13.6	17.9	14.9
MedAE	845.9	1088.2	1438.0	608.7	1360.3	482.6	580.3	436.8	870.0	471.3
Bias	-209.9	-1834.9	-1802.7	-841.9	5283.7	-166.2	140.1	-707.9	-1096.9	-76.5
Imprecision	2205.9	2687.5	2652.5	2361.4	8800.5	1495.3	1558.4	1653.4	2057.0	1194.2

Table A.6. The percentage of spatial units (districts and wards) for which the adjusted census population totals are within the 95% CIs of the modelled population estimates, calculated from the expanded and conservative subsets. Values are shown for all spatial units nationally, as well as with stratification by province.

		District-level		Ward-level	
		Expanded subset	Conservative subset	Expanded subset	Conservative subset
National		53.4%	75.0%	46.6%	69.5%
Province	Central	45.5%	90.9%	42.8%	80.0%
	Copperbelt	40.0%	40.0%	41.1%	59.7%
	Eastern	73.3%	60.0%	71.8%	66.5%
	Luapula	75.0%	83.3%	71.6%	78.1%
	Lusaka	16.7%	66.7%	28.8%	60.2%
	Muchinga	50.0%	100.0%	31.7%	85.6%
	Northern	50.0%	91.7%	42.8%	82.8%
	Northwestern	81.8%	72.7%	62.3%	68.3%
	Southern	33.3%	66.7%	33.5%	57.9%
	Western	50.0%	81.2%	36.1%	69.6%

Table A.7. The percentage of spatial units (districts and wards) for which the adjusted census population counts are within the 95% CIs of the modelled population estimates, calculated from the conservative subset. For those spatial units in which the adjusted census counts are outside the 95% CIs of the modelled estimates, the percentage of spatial units in which the lower 95% CI exceeds the adjusted census count, and the upper 95% CI is less than the adjusted census count is also shown. Values are shown for all spatial units nationally, as well as with stratification by province.

		Conservative subset - Districts			Conservative subset - Wards		
		within 95% CIs	L95% CI > census	U95% CI < census	within 95% CIs	L95% CI > census	U95% CI < census
National		75.0	4.3	20.7	69.5	6.8	23.7
Province	Central	90.9	0.0	9.1	80.0	5.6	14.4
	Copperbelt	40.0	0.0	60.0	59.7	2.4	37.9
	Eastern	60.0	0.0	40.0	66.5	1.0	32.5
	Luapula	83.3	0.0	16.7	78.1	3.6	18.3
	Lusaka	66.7	33.3	0.0	60.2	30.5	9.3
	Muchinga	100.0	0.0	0.0	85.6	1.9	12.5
	Northern	91.7	8.3	0.0	82.8	5.6	11.7
	Northwestern	72.7	0.0	27.3	68.3	2.4	29.3
	Southern	66.7	0.0	33.3	57.9	4.3	37.8
	Western	81.3	12.5	6.3	69.6	17.2	13.2

Table A.8. Summary statistics for each province, and nationally, concerning the count of grid cells with (i) modelled population estimates, (ii) census households, and (iii) census households but no modelled population estimates, alongside the (iv) census population totals, (v) the estimated proportion of the census population resident in grid cells without modelled population estimates, and (vi) the resulting adjusted census population estimates.

		Count of grid cells, with:			Census population		
		Modelled pop. estimates	Census HHs	Census HHs, no modelled estimates	Census population total	Proportion in grid cells with only HHs	Adjusted census population
National		2,273,498	1,525,866	380,754	19,693,423	0.12	17,382,074
Province	Central	430,585	259,959	72,709	2,261,336	0.19	1,827,265
	Copperbelt	212,130	130,533	29,627	2,768,192	0.06	2,603,311
	Eastern	268,300	187,684	41,711	2,462,682	0.11	2,193,971
	Luapula	137,669	99,822	20,031	1,519,478	0.09	1,378,652
	Lusaka	155,884	108,296	17,315	3,093,617	0.03	2,998,825
	Muchinga	155,608	102,214	28,670	922,213	0.18	758,693
	Northern	225,864	155,211	36,155	1,623,853	0.13	1,408,399
	Northwestern	132,942	96,792	24,252	1,278,357	0.13	1,111,249
	Southern	379,634	264,035	79,370	2,388,091	0.20	1,919,471
	Western	174,882	121,320	30,914	1,375,604	0.14	1,182,237

Dataset S1 (separate file).

The modelled population estimates and census population counts are provided as supporting datasets in an Excel spreadsheet (**ZMB_Modelled_Population_and_2022_Census.xlsx**). This is structured with one tab per administrative unit level (ward, district, province and national), and includes modelled estimates based on the conservative and expanded grid cell subsets, with modelled mean values and upper and lower 95% credible intervals (CIs). The variables included in the dataset are described in the following text:

- *PROVINCE_NAME* – name of the province in which the administrative unit is located
- *DISTRICT_NAME* – name of the district in which the administrative unit is located
- *CONSTITUENCY_NAME* – name of the constituency in which the administrative unit is located
- *WARD_NAME* – name of the ward in which the administrative unit is located
- *level* – administrative unit type (national, province, district, ward)
- *census22_population* – the total population for the administrative unit, as published in the 2022 Census results (Zambia Statistics Agency, 2024)
- *Pop_mean_all* – mean value of the modelled population estimates for the administrative unit, calculated from the posterior distribution for all grid cells with centroids within the administrative unit
- *Pop_l95_all* – lower 95% credible interval of the modelled population estimates for the administrative unit, calculated from the posterior distribution for all grid cells with centroids within the administrative unit
- *Pop_u95_all* – upper 95% credible interval of the modelled population estimates for the administrative unit, calculated from the posterior distribution for all grid cells with centroids within the administrative unit
- *Pop_mean_expSS* – mean value of the modelled population estimates for the administrative unit, calculated from the posterior distribution for the expanded subset of grid cells with centroids within the administrative unit
- *Pop_l95_expSS* – lower 95% credible interval of the modelled population estimates for the administrative unit, calculated from the posterior distribution for the expanded subset of grid cells with centroids within the administrative unit
- *Pop_u95_expSS* – upper 95% credible interval of the modelled population estimates for the administrative unit, calculated from the posterior distribution for the expanded subset of grid cells with centroids within the administrative unit
- *Pop_mean_conSS* – mean value of the modelled population estimates for the administrative unit, calculated from the posterior distribution for the conservative subset of grid cells with centroids within the administrative unit
- *Pop_l95_conSS* – lower 95% credible interval of the modelled population estimates for the administrative unit, calculated from the posterior distribution for the conservative subset of grid cells with centroids within the administrative unit
- *Pop_u95_conSS* – upper 95% credible interval of the modelled population estimates for the administrative unit, calculated from the posterior distribution for the conservative subset of grid cells with centroids within the administrative unit

SI References

Dooley, C., Chamberlain, H., Leasure, D., Membele, G., Lazar, A., & Tatem, A. (2021). Description of methods for the Zambia modelled population estimates from multiple routinely collected and geolocated survey data, version 1.0. <https://doi.org/10.5258/SOTON/WP00707>

WorldPop and Zambia Statistics Agency. (2024). Census disaggregated gridded population estimates for Zambia, based on the 2022 Census of Population and Housing preliminary results, version 2.0. <https://doi.org/https://doi.org/10.5258/SOTON/WP00783>

WorldPop. (2020). Bottom-up gridded population estimates for Zambia, version 1.0. <https://doi.org/10.5258/SOTON/WP00662>

Zambia Statistics Agency. (2024). Zambia 2022 Census of Population and Housing Summary Report: Volume Two (Revised). <https://www.zamstats.gov.zm/Publications/Revised%202022%20Census%20of%20Population%20and%20Housing%20Summary%20Report%20Volume%20II.pdf>