Release Statement

Spatio-temporally harmonised datasets for Democratic Republic of the Congo. Version 1.0

26 October 2023

This work is part of the WorldPop Global Demographic Data project with funding from the Bill and Melinda Gates Foundation (BMGF). Method, coding, and data production have been carried out by Thea Woods, Tom McKeen, Alexander Cunningham, Rhorom Priyatikanto and oversight was provided by Maksym Bondarenko.

CITATION

T. Woods, T. McKeen, A. Cunningham, R. Priyatikanto, A. Sorichetta , AJ Tatem and M. Bondarenko. 2023 *Spatio-temporally harmonised datasets for Democratic Republic of the Congo. Version 1.0*. University of Southampton: Southampton, UK. DOI: 10.5258/SOTON/WP00768.

RELEASE CONTENT

1. **COD_L0_GL2_v1.tif**
2. **COD_total_area_*{PIB|BCB}_**{gl|ms}_bf_GL2_v1.tif**
3. **COD_mean_area_*{PIB|BCB}_**{gl|ms}_bf_GL2_v1.tif**
4. **COD_cv_area_*{PIB|BCB}_**{gl|ms}_bf_GL2_v1.tif**
5. **COD_total_length_*{PIB|BCB}_**{gl|ms}_bf_GL2_v1.tif**
6. **COD_mean_length_*{PIB|BCB}_**{gl|ms}_bf_GL2_v1.tif**
7. **COD_cv_length_*{PIB|BCB}_**{gl|ms}_bf_GL2_v1.tif**
8. **COD_count_*{PIB|BCB}_**{gl|ms}_bf_GL2_v1.tif**
9. **COD_density_*{PIB|BCB}_**{gl|ms}_bf_GL2_v1.tif**
10. **COD_C3S-LC-L4-LCCS-***XXX_****YYY_100m_dist.tif**
11. **COD_waterbodies_dist_osm_GL2_v1.tif**
12. **COD_intersections_dist_osm_GL2_v1.tif**
13. **COD_highway_dist_osm_GL2_v1.tif**
14. **COD_WDPA_dist-***XXX _cat1_100m.tif**
15. **COD_VNL_v21_npp-***XXX _average_masked_bilE.tif**
16. **COD_count_ms_roads_GL2_v1.tif**
17. **COD_binary_ms_roads_GL2_v1.tif**
18. **COD_density_ms_roads_GL2_v1.tif**
19. **COD_dist_ms_roads_GL2_v1.tif**
20. **COD_total_length_ms_roads_GL2_v1.tif**
21. **COD_water_mask_GL2_v1.tif**
22. **COD_temperate_avg_GL2_v1.tif**
23. **COD_prec_GL2_v1.tif**
24. **COD_elevation_GL2_v1.tif**
25. **COD_slope_GL2_v1.tif**
26. **COD_costline_dist_GL2_v1.tif**

* Building metrics are calculated based on the pixel in which their respective centroids are located (**BCB** method), or the pixel(s) that their geometries intersect (Pixel intersected Based – **PIB** method). In-depth explanations of each method follow in the methods overview below.
** **gl -** google Open Buildings V3, **ms** – Microsoft Building Footprints 2023
*** **XXX** - year
******YYY -** class

FILE DESCRIPTIONS

The spatially referenced geotiff grid layers introduced in this description all have a resolution of 3-arc seconds (~100m at the equator) and have a geographic spatial reference (WGS84). All geotiffs have a NoData value of -99999 (pixels that do not contain any buildings).

COD_L0_GL2_v1.tif
Mastergrid for Democratic Republic of the Congo.

COD_count_*{PIB|BCB}_{gl|ms}_bf_GL2_v1.tif**
Each grid cell of this dataset represents the counts of buildings within the grids'/pixels' bounds based on the 2 methods employed.

COD_density_*{PIB|BCB}_{gl|ms}_bf_GL2_v1.tif**
Each grid cell of this dataset represents the count of buildings within respective grid cells, divided the grid cell's area.

COD_total_area_*{PIB|BCB}_{gl|ms}_bf_GL2_v1.tif**
Grid cells in this dataset represent the total building footprint area within respective grid cells. For buildings that straddle multiple grid cells, the grid cell in which the buildings' centroids are located will be allocated the area value. Note: For the **BCB** method, total building area may exceed the area of a grid cell if the centroid of a large building falls within the grid cell.

COD_mean_length_*{PIB|BCB}_{gl|ms}_bf_GL2_v1.tif**
Grid cells in this dataset represent the mean area of all buildings whose buildings are within the respective grid cells.

COD_cv_area_*{PIB|BCB}_{gl|ms}_bf_GL2_v1.tif**
Grid cells in this dataset represent a grid cell level coefficient of variation of building areas for all buildings inside a grid cell. Coefficient of variation is the standard deviation divided by the mean.

COD_total_length_*{PIB|BCB}_{gl|ms}_bf_GL2_v1.tif**
Total length per grid cell is therefore a sum of the perimeters of all the buildings inside a grid cell.

COD_mean_length_*{PIB|BCB}_{gl|ms}_bf_GL2_v1.tif**
Mean perimeter length of all the buildings inside a grid cell.

COD_cv_length_*{PIB|BCB}_{gl|ms}_bf_GL2_v1.tif**
Grid cells in this dataset represent the coefficient of variation of building lengths (perimeter) for all buildings inside a grid cell.

COD_C3S-LC-L4-LCCS-*XXX_****YYY_100m_dist.tif**
Distance to edges of reclassified ESA-CCI-LC classes 2015-2020. The values of the raster are the from the cell centres to the nearest feature.
Classes:

- 11 - Cropland, Natural Vegetation
- 40 - Tree Cover
- 130 - Shrubland
- 140 - Herbaceous Cover, Grassland, Mosses
- 150 - Sparse Vegetation
- 160 - Tree/Herbaceous Cover, Flooded, Fresh/Saline/Brackish Water
- 190 - Urban Areas
- 200 - Bare Areas
- 210 - Water Bodies, Permanent Snow And Ice

COD_waterbodies_dist_osm_GL2_v1.tif
Distance to OSM major waterways. OSM data up to 2023-01-17T21:21:52Z

COD_intersections_dist_osm_GL2_v1.tif
Distance to OSM major road intersections. OSM data up to 2023-01-17T21:21:52Z

COD_highway_dist_osm_GL2_v1.tif
Distance to OSM major roads. OSM data up to 2023-01-17T21:21:52Z

COD_WDPA_dist-*XXX _cat1_100m.tif**
Distance to IUCN strict nature reserve and wilderness area edges 2015-2022

COD_VNL_v21_npp-*XXX _average_masked_bilE.tif**
Night lights 2015-2021 VIIRS NTL 2.1

COD_count_ms_roads_GL2_v1.tif
Each grid cell of this dataset represents the counts of roads within the grids'/pixels'.
https://github.com/microsoft/RoadDetections

COD_binary_ms_roads_GL2_v1.tif
Grid cells in this dataset with value 1 indicate the presence of at road. https://github.com/microsoft/RoadDetections

COD_dist_ms_roads_GL2_v1.tif
Distance to roads. https://github.com/microsoft/RoadDetections

COD_water_mask_GL2_v1.tif
Each grid cell of this dataset represents the percentage of water within the grids'/pixels'. (Source: WorldCover 2021 v200 10m https://esa-worldcover.org/en/data-access)

COD_temperate_avg_GL2_v1.tif
Temperature . (Source: WorldClim 2.1, 1970-2000)

COD_prec_GL2_v1.tif
Precipitation. (Source: WorldClim 2.1, 1970-2000)

COD_elevation_GL2_v1.tif
Elevation

COD_slope_GL2_v1.tif
Slope

COD_costline_dist_GL2_v1.tif
Distance to open-water coastline .

Building centroid based and pixel intersected based method overview.

The **Building Centroid Based** (BCB) method allocates values to grid cells based on building centroids located within respective grid cells' bounds. In cases where buildings straddle multiple grid cells, their metrics (area, length, etc.) will only be allocated to the grid cell in which the buildings' centroids are located, as shown in Figure 1. All grid datasets are 3 arc second (approximately 100m at the equator) projected to WGS84. This method is more suited to building count and distance-to-neighbour related metrics.

The **Pixel intersected Based** (PIB) method makes use of the geometric intersections between building footprints and grid cells to calculate values, resulting in building metrics being allocated to all grid cells in which building footprints are present. Figure 1. All grid datasets are 3 arc second (approximately 100m at the equator) projected to WGS84. This method results in a more continuous grid, with exact measurements related to building area and length being allocated to the pixels in which they are proportionally located.

It should be noted that the number of valid (non-NoData) grid cells in the PIB method will most likely be greater than the corresponding dataset using the BCB method, due to the BCB method only considering building centroids in its

calculations. This may result in slightly different grid cell extents when comparing outputs from the respective methods. Please read above file descriptions to select the most appropriate dataset for your project's requirements.

Figure 1.

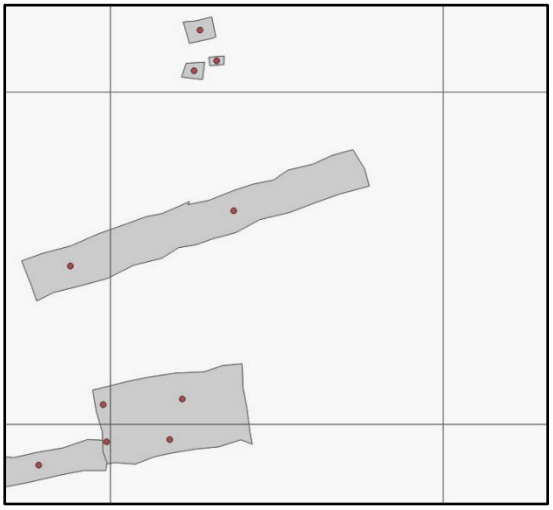

Figure 2.