## Release Statement

## Gridded population estimates for Ukraine using UN COD-PS estimates 2020, version 2.0

28 March 2022

These data were produced by WorldPop at the University of Southampton [1] and the 'Smart Cities and Spatial Development' team at the German Remote Sensing Data Center (DFD) of the German Aerospace Center (DLR) [2]. These data include gridded estimates of population at approximately 100m and 1km resolution for 2020, along with estimates of the number of people belonging to individual age-sex groups. These results were produced using subnational population estimates for Ukraine in 2020 provided in the Common Operational Dataset on Population Statistics (COD-PS) [6] and building height/area/fraction/volume covariates extracted from the World Settlement Footprint (WSF) imperviousness and WSF-3D by DLR [3,4]. The constrained top-down disaggregation method was used to produce the datasets. The modelling work and geospatial data processing was led by Bondarenko M., Palacios-Lopez D., Sorichetta A., Leasure D.R., ,Zeidler J., Marconcini M., and Esch T.. Oversight was provided by Tatem A.J. Internal WorldPop peer reviews that helped to improve the results and documentation was provided by Lazar A.N..



*The authors followed rigorous procedures designed to ensure that the used data, the applied method and thus the results are appropriate and of reasonable quality. If users encounter apparent errors or misstatements, they should contact WorldPop at release@worldpop.org.*



### CITATION

Bondarenko M., Palacios-Lopez D., Sorichetta A., Leasure D.R., Zeidler J., Marconcini, M., Esch T., and Tatem A.J. 2022 *Gridded population estimates for Ukraine using UN COD-PS estimates 2020, version 2.0*. WorldPop and DLR, University of Southampton. doi:10.5258/SOTON/WP00735.

### MAIN DATA SOURCES

- The German Aerospace Centre's (DLR) WSF imperviousness [3] and WSF 3D products (WSF-3D) [4].
- Subnational population estimates for Ukraine in 2020 provided in the Common Operational Dataset on Population Statistics (COD-PS) [6]. The subnational population estimates were produced using baseline information from the 2001 Population Census of Ukraine and annual birth and death registration data since then.
- Subnational Administrative Boundaries for Ukraine provided by OCHA [7].
- Geospatial covariate layers available at WorldPop [12].

## RELEASE CONTENT

1. ukr_pop_2020_100m_constrained_v2.zip
2. ukr_pop_2020_1km_constrained_v2.zip
3. ukr_agesex_2020_100m_constrained_v2.zip
4. ukr_agesex_2020_1km_constrained_v2.zip
5. ukr_agesex_0_18_2020_100m_constrained_v2.zip
6. ukr_agesex_0_18_2020_1km_constrained_v2.zip

## FILE DESCRIPTIONS

The projection for all GIS files is the geographic coordinate system WGS84 (World Geodetic System 1984).

**ukr_pop_2020_100m_constrained_v2.zip**

This geotiff raster, at a spatial resolution of 3 arc-seconds (approximately 100m at the equator), contains estimates of total population size per grid cell across Ukraine. NA values represent areas that were mapped as unsettled based on the DLR settlement layer [3]. These data are stored as floating-point numbers rather than integers to avoid rounding errors in aggregated populations for larger areas.

**ukr_pop_2020_1km_constrained_v2.zip**

This geotiff raster, at a spatial resolution of 30 arc-seconds (approximately 1km at the equator), contains estimates of total population size per grid cell across Ukraine. NA values represent areas that were mapped as unsettled based on the DLR settlement layer [5]. The dataset was produced by aggregating the *ukr_pop_2020_100m_constrained_v2.tif* dataset to 1km.

**ukr_agesex_2020_100m_constrained_v2.zip**

This zip file contains 34 raster files in geotiff format at a spatial resolution of 3 arc-seconds (approximately 100m at the equator). Each raster provides gridded population estimates for an age-sex group of settled areas (NA represent unsettled areas). Files are labelled with either an “M” (male) or an “F” (female) followed by the age-range of the group (five year bins). For instance, “F_00_04” and “M_05_09” are population counts of under 5 year olds for females and between 5 and 9 years old for males, respectively. Eighty year olds and over are represented in the groups “F_80Plus” and “M_80Plus”. These data were produced using age-sex national proportions from COD-PS [6]. The age-sex proportions were applied to the gridded population estimates (*ukr_pop_2020_100m_constrained_v2*) to allocate the population to the different age-sex classes. While this data represents population counts, values contain decimals, i.e. fractions of people. This is because we do not estimate which grid cell each individual in a given age group occupies. For this reason, it is advised to aggregate the rasters at a coarser scale. For example, if four grid cells next to each other have values of 0.25 this indicates that there is 1 person of that age group somewhere in those four grid cells.

**ukr_agesex_2020_1km_constrained_v2.zip**

This zip file contains 34 rasters in geotiff format at a spatial resolution of 30 arc-seconds (approximately 1km at the equator). Each raster provides gridded population estimates for an age-sex group of settled areas (NA represent unsettled areas). These datasets were produced by aggregating the *ukr_agesex_2020_100m_constrained_v2* datasets respectively to 1km.

**ukr_agesex_0_18_2020_100m_constrained_v2.zip**

This zip file contains 3 rasters in geotiff format at a spatial resolution of 3 arc-seconds (approximately 100m at the equator). Each raster provides gridded population estimates for children-age (0 to 18 years old) of settled areas (NA represent unsettled areas). Files are labelled with either an “M” (male) or an “F” (female) or “F_M” for both male and female. These data were created by applying the Sprague multipliers to “*ukr_agesex_2020_100m_constrained_v2* “ raster data. Sprague multipliers are used to interpolate data

and obtain population estimates by single years of age and these are then aggregated to 0 to 18 years of age.

**ukr_agesex_0_18_2020_1km_constrained_v2.zip**
This zip file contains 3 rasters in geotiff format at a spatial resolution of 30 arc-seconds (approximately 1km at the equator). Each raster provides gridded population estimates for children-age (0 to 18 years old) of settled areas (NA represent unsettled areas). Files are labelled with either an "M" (male) or an "F" (female) or "F_M" for both male and female. These data were created by applying Sprague multipliers to "*ukr_agesex_2020_1km_constrained_v2* " raster data. The Sprague multipliers are used to interpolate data and obtain population estimates by single years of age and these are then aggregated to 0 to 18 years of age.

## RELEASE HISTORY

- Version 2.0 (28 March 2022) [https://dx.doi.org/10.5258/SOTON/WP00735]
    - Refinement of gridded population estimates using more recent settlement data based on the German Aerospace Centre's (DLR) World Settlement Footprint 3D product (WSF-3D) [3,4].
- Version 1.0 (14 March 2022) [https://dx.doi.org/10.5258/SOTON/WP00734] Original release of Ukraine 2020 population dataset

## METHODS

**Pre-processing:** Subnational Administrative Boundaries provided by OCHA [7] were nibbled (i.e. cells with no data are replaced with the values of the nearest neighbors) to match the WorldPop mastergrid and avoid mismatch with the WorldPop covariates [12].

**Modelling**: Building height/area/fraction/volume per pixel were extracted from satellite data by DLR in addition to classifying pixels as residential or non-residential following the methodology presented in [5], in which spatial metrics derived solely from the WSF3D datasets are used to train a Random Forest classifier. Comparably, the population modelling described below used the Random Forest (RF)-based dasymetric mapping approach (Stevens et al., 2015 [10]) implemented in the popRF 'R' package [9] based on the Breiman (2001) [11] algorithm.

The UN COD-PS [6] population projections do not have sufficient admin units to apply the RF methodology well. Therefore admin 3 projected population data from WorldPop [12] as used to train the RF model and produce an unconstrained prediction weighting layer, where all non-residential pixels were given a value of zero in the building covariate inputs (APPENDIX 2). The model could explain 91.7% of the population input variance. The unsettled areas of this unconstrained weighting layer were then removed by using the DLR settlement layer containing all buildings (residential and non-residential) as a mask, thus creating a constrained weighting layer. Finally, this constrained prediction weighting layer was used for dasymetric redistribution of the UN COD-PS [6] population estimates 2020 (APPENDIX 1). The gridded population estimates were then combined with the COD-PS [6] age/sex pyramid table for Ukraine to produce gridded population estimates for females and males at regular age intervals.

**WORK CITED**


1. WorldPop, University of Southampton, UK https://www.worldpop.org

2. German Aerospace Center (DLR), Wessling, Germany https://www.dlr.de

3. Marconcini, M., Metz-Marconcini, A., Esch, T., & Gorelick, N. (2021). Understanding Current Trends in Global Urbanisation-The World Settlement Footprint Suite.*GI-Forum 2021 (1)* https://doi.org/ 10.1553/giscience2021_01_s33

4. Esch, T., Brzoska, E., Dech, S., Leutner, B., Palacios-Lopez, D., Metz-Marconcini, A., Marconcini, M., Roth, A., Zeidler, J. (2022): World Settlement Footprint 3D - A first three-dimensional survey of the global building stock. Remote Sensing of Environment, Volume 270, 2022, 112877, ISSN 0034-4257. https://doi.org/10.1016/j.rse.2021.112877

5. Palacios-Lopez, D., Esch, T., MacManus, K., Marconcini, M., Sorichetta, A., Yetman, G., Zeidler, J., Dech, S., Tatem, A.J., Reinartz, P. (2022): Towards an Improved Large-Scale Gridded Population Dataset: A Pan-European Study on the Integration of 3D Settlement Data into Population Modelling. Remote Sens. 2022, 14, 325. https://doi.org/10.3390/rs14020325

6. UNFPA Ukraine - Subnational Population Statistics https://data.humdata.org/dataset/cod-ps-ukr

7. OCHA Ukraine - Subnational Administrative Boundaries https://data.humdata.org/dataset/cod-ab-ukr

8. G. Calot and J.-P. Sardon. Methodology for the calculation of Eurostat's demographic indicators. Detailed report by the European Demographic Observatory. Brussels: European Commission (Population and Social Conditions 3/2003/F/no 26).

9. Bondarenko M., Nieves J.J., Forrest R.S., Andrea E.G., Jochem C., Kerr D., and Sorichetta A. (2021): popRF: Random Forest-informed Population Disaggregation R package, _Comprehensive R Archive Network (CRAN)_,url:https://cran.rproject.org/package=popRF.

10. Stevens FR, Gaughan AE, Linard C, Tatem AJ (2015) Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data. PLoS ONE 10(2): e0107042. https://doi.org/10.1371/journal.pone.0107042

11. Breiman L. Random Forests. Mach Learn. 2001 [cited 2013 Jan 24].45(1):5–32. Available: http://link.springer.com/article/10.1023/A:1010933404324

12. WorldPop (www.worldpop.org - School of Geography and Environmental Science, University of Southampton; Department of Geography and Geosciences, University of Louisville; Departement de Geographie, Universite de Namur) and Center for International Earth Science Information Network (CIESIN), Columbia University (2018). Global High Resolution Population Denominators

Project - Funded by the Bill and Melinda Gates Foundation (OPP1134076).https://dx.doi.org/10.5258/SOTON/WP00644

13. WorldPop (www.worldpop.org - School of Geography and Environmental Science, University of Southampton; Department of Geography and Geosciences, University of Louisville; Departement de Geographie, Universite de Namur) and Center for International Earth Science Information Network (CIESIN), Columbia University. 2018b. Global High Resolution Population Denominators Project - Funded by The Bill and Melinda Gates Foundation (OPP1134076). Available from ftp://ftp.worldpop.org/GIS/Population/Global_2000_2020/CensusTables/

**APPENDIX 1**

Table of *ADM1_PCODE* Subnational Administrative Boundaries for Ukraine provided by OCHA. Source https://data.humdata.org/dataset/cod-ps-ukr

| ISO3 | ADM1_NAME | ADM1_COD_AB_EN | ADM1_PCODE |
|---|---|---|---|
| UKR | Cherkasy oblast | Cherkaska | UA71 |
| UKR | Chernihiv oblast | Chernihivska | UA74 |
| UKR | Chernivtsi oblast | Chernivetska | UA73 |
| UKR | Dnipropetrovsk oblast | Dnipropetrovska | UA12 |
| UKR | Donetsk oblast | Donetska | UA14 |
| UKR | Ivano-Frankivsk oblast | Ivano-Frankivska | UA26 |
| UKR | Kharkiv oblast | Kharkivska | UA63 |
| UKR | Kherson oblast | Khersonska | UA65 |
| UKR | Khmelnytskiy oblast | Khmelnytska | UA68 |
| UKR | Kirovohrad oblast | Kirovohradska | UA35 |
| UKR | Kyiv city | Kyivska | UA80 |
| UKR | Kyiv oblast | Kyivska | UA32 |
| UKR | Luhansk oblast | Luhanska | UA44 |
| UKR | Lviv oblast | Lvivska | UA46 |
| UKR | Mykolayiv oblast | Mykolaivska | UA48 |
| UKR | Odesa oblast | Odeska | UA51 |
| UKR | Poltava oblast | Poltavska | UA53 |
| UKR | Rivne oblast | Rivnenska | UA56 |
| UKR | Sumy oblast | Sumska | UA59 |
| UKR | Ternopil oblast | Ternopilska | UA61 |
| UKR | Vinnytsya oblast | Vinnytska | UA05 |
| UKR | Volyn oblast | Volynska | UA07 |
| UKR | Zakarpattya oblast | Zakarpatska | UA21 |
| UKR | Zaporizhzhya oblast | Zaporizka | UA23 |
| UKR | Zhytomyr oblast | Zhytomyrska | UA18 |
| UKR | Autonomous Republic Crimea | Avtonomna Respublika Krym | UA01 |
| UKR | Sevastopol | Sevastopilska | UA85 |

Table of *adm4_PCODE* Subnational Administrative Boundaries for Ukraine provided by OCHA. Source https://data.humdata.org/dataset/cod-ps-ukr

| ISO3 | ADM1_NAME | City_Name_EN | adm4_PCODE |
|---|---|---|---|
| UKR | Cherkasy oblast Cherkasy city | Cherkasy | UA7108049001 |
| UKR | Chernihiv oblast Chernihiv city | Chernihiv | UA7410039001 |
| UKR | Chernivtsi oblast Chernivtsi city | Chernivtsi | UA7306061001 |
| UKR | Dnipropetrovsk oblast Kamianske city | Kamianske | UA1204015001 |
| UKR | Dnipropetrovsk oblast Kryvyi Rih city | Kryvyi Rih | UA1206017001 |
| UKR | Dnipropetrovsk oblast Nikopol city | Nikopol | UA1208005001 |
| UKR | Dnipropetrovsk oblast Pavlohrad city | Pavlohrad | UA1212007001 |
| UKR | Dnipropetrovsk oblast Dnipro | Dnipro | UA1202001001 |
| UKR | Ivano-Frankivsk oblast | Ivano-Frankivsk | UA2604019001 |
| UKR | Kharkiv oblast Kharkiv city | Kharkiv | UA6312027001 |
| UKR | Kherson oblast Kherson city | Kherson | UA6510015001 |
| UKR | Khmelnytskiy oblast Khmelnytskiy city | Khmelnytskiy | UA6804047001 |
| UKR | Kirovohrad oblast Kropyvnytskyi city | Kropyvnytskyi | UA3504021001 |
| UKR | Kyiv oblast Bila Tserkva city | Bila Tserkva | UA3202001001 |
| UKR | Kyiv oblast Brovary city | Brovary | UA3206005001 |
| UKR | Lviv oblast Lviv city | Lviv | UA4606025001 |
| UKR | Mykolayiv oblast Mykolayiv city | Mykolayiv | UA4806015001 |
| UKR | Odesa oblast Odesa city | Odesa | UA5110027001 |
| UKR | Poltava oblast Kremenchuk city | Kremenchuk | UA5302011001 |
| UKR | Poltava oblast Poltava city | Poltava | UA5308037001 |
| UKR | Rivne oblast Rivne city | Rivne | UA5606047001 |
| UKR | Sumy oblast Sumy city | Sumy | UA5908027001 |
| UKR | Ternopil oblast Ternopil city | Ternopil | UA6104049001 |
| UKR | Vinnytsya oblast Vinnytsya city | Vinnytsya | UA0502003001 |
| UKR | Volyn oblasts Lutsk city | Lutsk | UA0708017001 |
| UKR | Zakarpattya oblast Uzhhorod city | Uzhhorod | UA2110023001 |
| UKR | Zaporizhzhya oblast Melitopol city | Melitopol | UA2308007001 |
| UKR | Zaporizhzhya oblast | Zaporizhzhya | UA2306007001 |
| UKR | Zaporizhzhya oblasts  Berdiansk city | Berdiansk | UA2302005001 |
| UKR | Zhytomyr oblast Zhytomyr city | Zhytomyr | UA1804019001 |

**APPENDIX 2**

List of covariates.

| Name of the covariate | Description |
|---|---|
| ukr_grid_100m_ccilc_dst011_2015.tif | Distance to ESA-CCI-LC cultivated area edges 2015 |
| ukr_grid_100m_ccilc_dst040_2015.tif | Distance to ESA-CCI-LC woody-tree area edges 2015 |
| ukr_grid_100m_ccilc_dst130_2015.tif | Distance to ESA-CCI-LC shrub area edges 2015 |

| | |
|---|---|
| ukr_grid_100m_ccilc_dst140_2015.tif | Distance to ESA-CCI-LC herbaceous area edges 2015 |
| ukr_grid_100m_ccilc_dst150_2015.tif | Distance to ESA-CCI-LC sparse vegetation area edges 2015 |
| ukr_grid_100m_ccilc_dst160_2015.tif | Distance to ESA-CCI-LC aquatic vegetation area edges 2015 |
| ukr_grid_100m_ccilc_dst190_2015.tif | Distance to ESA-CCI-LC artificial surface edges 2015 |
| ukr_grid_100m_ccilc_dst200_2015.tif | Distance to ESA-CCI-LC bare area edges 2015 |
| ukr_grid_100m_cciwat_dst.tif | ESA-CCI-LC inland waterbodies 2000-2012 |
| ukr_grid_100m_gpw4coast_dst.tif | Distance to coastline 2000-2020 |
| ukr_grid_100m_osmint_dst.tif | Distance to OSM major road intersections |
| ukr_grid_100m_osmriv_dst.tif | Distance to OSM major waterways |
| ukr_grid_100m_osmroa_dst.tif | Distance to OSM major roads |
| ukr_grid_100m_px_area.tif | Grid-cell surface areas |
| ukr_grid_100m_slope.tif | SRTM-based slope 2000 ( SRTM is Shuttle Radar Topography Mission) |
| ukr_grid_100m_topo.tif | SRTM elevation 2000 |
| ukr_grid_100m_viirs_2016.tif | VIIRS night-time lights 2016 (VIIRS is Visible Infrared Imaging Radiometer Suite ) |
| ukr_grid_100m_wclim_prec.tif | Current average annual total precipitation |
| ukr_grid_100m_wclim_temp.tif | Current average annual temperature |
| ukr_grid_100m_wdpa_cat1_dst_2017.tif | Distance to IUCN strict nature reserve and wilderness area edges 2017 |
| LandScanHDUkraine_dst.tif | Distance to settlement ORNL LandScan High Definition (HD) Data for Ukraine |
| Building_Height.tif | DLR WSF3D building height |
| Building_Area.tif | DLR WSF3D building area |
| Building_Fraction | DLR WSF3D building fraction |
| Building_Volume | DLR WSF3D building volume |