Release Statement

Gridded population estimates for Turkey using UN COD-PS estimates 2021, version 1.0

22 February 2023

These data were produced by WorldPop at the University of Southampton. These data include gridded estimates of population at approximately 100m for 2021, along with estimates of the number of people belonging to individual age-sex groups. These results were produced using subnational population estimates for Turkey in 2021 provided in the Common Operational Dataset on Population Statistics (COD-PS) and built-up surfaces/volumes/height covariates extracted from GHSL datasets. The constrained and unconstrained top-down disaggregation method was used to produce the datasets. The modelling work and geospatial data processing was led by Bondarenko M., Priyatikanto R., Sorichetta A. Oversight was provided by Tatem A.J.

These data may be distributed using a Creative Commons Attribution 4.0 International (CC BY 4.0) License, specified in legal code. Contact release@worldpop.org for more information.

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.

WorldPop, University of Southampton, and their sponsors offer these data on a "where is, as is" basis; do not offer an express or implied warranty of any kind; do not guarantee the quality, applicability, accuracy, reliability or completeness of any data provided; and shall not be liable for incidental, consequential, or special damages arising out of the use of any data that they offer.

CITATION
Bondarenko M., Priyatikanto R., Sorichetta A., and Tatem A.J. 2022 *Gridded population estimates for Turkey using UN COD-PS estimates 2021, version 1.0*. WorldPop, University of Southampton. doi:10.5258/SOTON/WP00758.

MAIN DATA SOURCES

- GHS-BUILT-S R2022A - GHS built-up surface grid, derived from Sentinel-2 composite and Landsat, multitemporal (1975-2030) [1].
- Subnational population estimates for Turkey in 2021 provided in the Common Operational Dataset on Population Statistics (COD-PS). It is an update to an earlier COD-PS released by UNFPA [2].
- Subnational Administrative Boundaries for Turkey provided by Regional IM Working Group - Europe [3].
- Geospatial covariate layers available at WorldPop [4].
- OpenStreetMap [5] [Data file from 2023-01-17]
- WorldCover 10m 2021 v200 [6]
- VIIRS Nighttime Day/Night Band Composites Version 1 [7]. 2022 time series was extracted from Image Collections "NOAA/VIIRS/DNB/MONTHLY_V1/VCMCFG" using Google Earth Engine (GEE) platform.
- TerraClimate: Monthly Climate and Climatic Water Balance for Global Terrestrial Surfaces, University of Idaho [8]. 2022 time series was extracted from Image Collections "IDAHO_EPSCOR/TERRACLIMATE" using GEE platform.

- MOD11A2.061 Terra Land Surface Temperature and Emissivity 8-Day Global 1km [9]. 2022 time series was extracted from Image Collections "MODIS/061/MOD11A2" using GEE platform.

RELEASE CONTENT
1. tur_pop_2021_100m_constrained_v1.zip
2. tur_pop_2021_100m_unconstrained_v1.zip
3. tur_pop_2021_1km_constrained_v1.zip
4. tur_pop_2021_1km_unconstrained_v1.zip
5. tur_agesex_2021_100m_constrained_v1.zip
6. tur_agesex_2021_100m_unconstrained_v1.zip
7. tur_agesex_2021_1km_constrained_v1.zip
8. tur_agesex_2021_1km_unconstrained_v1.zip

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

tur_pop_2021_100m_constrained_v1.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 Turkey. NA values represent areas that were mapped as unsettled based on the GHS settlement layer [1]. These data are stored as floating-point numbers rather than integers to avoid rounding errors in aggregated populations for larger areas.

tur_pop_2021_100m_unconstrained_v1.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 Turkey. NA values represent areas that were mapped as water. These data are stored as floating-point numbers rather than integers to avoid rounding errors in aggregated populations for larger areas.

tur_pop_2021_1km_constrained_v1.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 Turkey. NA values represent areas that were mapped as unsettled based on the GHS settlement layer [1]. The dataset was produced by aggregating *tur_pop_2021_100m_constrained_v1.tif* dataset to 1km.

tur_pop_2021_1km_unconstrained_v1.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 Turkey. NA values represent areas that were mapped as water. The dataset was produced by aggregating *tur_pop_2021_100m_unconstrained_v1_0.tif* dataset to 1km.

tur_agesex_2021_100m_constrained_v1.zip
This zip file contains 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 [2]. The age-sex proportions were applied to the gridded

population estimates (*tur_pop_2021_100m_constrained_v1*) 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.

tur_agesex_2021_100m_uconstrained_v1.zip
This zip file contains 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 surface waters). 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 [2]. The age-sex proportions were applied to the gridded population estimates (*tur_pop_2021_100m_uconstrained_v1*) 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.

tur_agesex_2021_1km_constrained_v1.zip
This zip file contains 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 *tur_agesex_2021_100m_constrained_v1_0* datasets respectively to 1km

tur_agesex_2021_1km_uconstrained _v1.zip
This zip file contains raster files 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 across Turkey (NA represent surface waters). These datasets were produced by aggregating *tur_agesex_2021_100m_uconstrained_v1* datasets respectively to 1km.

METHODS

Pre-processing: Subnational Administrative Boundaries provided by OCHA 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.

Modelling: Building area/volume per pixel were extracted from GHS in addition to classifying pixels as residential or non-residential. 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 [11] based on the Breiman (2001) [12] algorithm.

The UN COD-PS [2] population was 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 built-up covariate inputs (APPENDIX 1). The model could explain 96.1% of the population input variance. The unsettled areas of this unconstrained weighting layer were then removed by using the GHS settlement layer containing all built-up surfaces (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 [2] population estimates 2021. The gridded population estimates were then combined with the COD-PS [2] age/sex pyramid table for Turkey to produce gridded population estimates for females and males at regular age intervals.

WORK CITED

1. Pesaresi, Martino; Politis, Panagiotis (2022): GHS-BUILT-S R2022A - GHS built-up surface grid, derived from Sentinel-2 composite and Landsat, multitemporal (1975-2030). European Commission, Joint Research Centre (JRC) [Dataset] DOI:10.2905/D07D81B4-7680-4D28-B896-583745C27085 PID:http://data.europa.eu/89h/d07d81b4-7680-4d28-b896-583745c27085

2. UNFPA Türkiye - Subnational Population Statistics https://data.humdata.org/dataset/cod-ps-tur

3. Regional IM Working Group - Europe - Subnational Administrative Boundaries https://data.humdata.org/dataset/cod-ab-tur

4. 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

5. OpenStreetMap contributors. (2015) Planet dump [Data file from 2023-01-17]. Retrieved from https://download.geofabrik.de/.

6. Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S., Lesiv, M., Herold, M., Tsendbazar, N.E., Xu, P., Ramoino, F., Arino, O., 2022. ESA WorldCover 10 m 2021 v200. https://doi.org/10.5281/zenodo.7254221

7. C. D. Elvidge, K. Baugh, M. Zhizhin, F. C. Hsu, and T. Ghosh, "VIIRS night-time lights," International Journal of Remote Sensing, vol. 38, pp. 5860–5879, 2017

8. Abatzoglou, J.T., S.Z. Dobrowski, S.A. Parks, K.C. Hegewisch, 2018, Terraclimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958-2015, Scientific Data 5:170191, doi:10.1038/sdata.2017.191

9. ASTER Mount Gariwang image from 2018 was retrieved on 2023-02-01 from https://lpdaac.usgs.gov, maintained by the NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC) at the USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota. 2018, https://lpdaac.usgs.gov/resources/data-action/aster-ultimate-2018-winter-olympics-observer/.

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. 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.

12. 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

APPENDIX 2

List of covariates.

Name of the covariate	Description
tur_grid_100m_ccilc_dst010_2021.tif	Distance to ESA-CCI-LC Tree cover 2021
tur_grid_100m_ccilc_dst020_2021.tif	Distance to ESA-CCI-LC Shrubland 2021
tur_grid_100m_ccilc_dst030_2021.tif	Distance to ESA-CCI-LC Grassland 2021
tur_grid_100m_ccilc_dst040_2021.tif	Distance to ESA-CCI-LC Cropland 2021
tur_grid_100m_ccilc_dst060_2021.tif	Distance to ESA-CCI-LC Bare/ sparse vegetation 2021
tur_grid_100m_ccilc_dst080_2021.tif	Distance to ESA-CCI-LC Permanent water bodies 2021
tur_grid_100m_gpw4coast_dst.tif	Distance to coastline 2000-2020
tur_grid_100m_osmint_dst_2022.tif	Distance to OSM major road intersections 2022
tur_grid_100m_osmriv_dst_2022.tif	Distance to OSM major waterways 2022
tur_grid_100m_osmroa_dst_2022.tif	Distance to OSM major roads 2022
tur_grid_100m_px_area.tif	Grid-cell surface areas
tur_grid_100m_slope.tif	SRTM-based slope 2000 (SRTM is Shuttle Radar Topography Mission)
tur_grid_100m_topo.tif	SRTM elevation 2000
tur_grid_100m_nighttime_2022.tif	Night-time lights 2022
tur_grid_100m_prec.tif	Current average annual total precipitation 2022
tur_grid_100m_temperature.tif	Current average annual temperature 2022
tur_grid_100m_wdpa_cat1_dst_2017.tif	Distance to IUCN strict nature reserve and wilderness area edges 2017
BUILTUP_dist.tif	Distance to settlement GHSL 2021
BUILT_S_2021.tif	GHSL built surfaces 2021
BUILT_V_2021.tif	GHSL built volume 2021