# Technical documentation for the hybrid 100-m global land cover dataset with Local Climate Zones for WRF

Matthias Demuzere[1], Cenlin He[2*], Alberto Martilli[3], Andrea Zonato[4,5]

[1] Ruhr-Universität Bochum, Bochum, Germany
[2] National Center for Atmospheric Research (NCAR), Boulder, CO, USA
[3] Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain
[4] University of Trento, Trento, Italy
[5] Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands

Date: 27 February 2023
*Corresponding author: Cenlin He (cenlinhe@ucar.edu)

---

## Summary

This work describes the development of a hybrid 100-m global land cover dataset and its implementation in the state-of-the-art Weather Research and Forecasting (WRF) model version 4.5. This hybrid CGLC-MODIS-LCZ dataset is based on 1) the Copernicus Global Land Service Land Cover (CGLC) product resampled to MODIS IGBP classes (CGLC-MODIS), and 2) the global map of Local Climate Zones (LCZ) that describes the heterogeneous urban land surface. Both the CGLC and LCZ products are available at a 100-m spatial resolution, are representative for the year 2018, and cover -180°W to 180°E and -60°S to 78°N. Remaining areas are filled with the MODIS land cover classes. This dataset has been implemented into the WRF Preprocessing System (WPS) as tiled binary data files[1] with a new GEOGRID table entry[2] to allow WRF/WPS users to flexibly use this dataset in their studies particularly relevant for urban modeling applications.

## 1. Introduction

Since the pioneering work of Brousse et al. (2016) and Martilli et al. (2016), the World Urban Database and Access Portal Tools (WUDAPT, Ching et al., 2018) level-0 Local Climate Zone (LCZ, Stewart and Oke, 2012) maps have been used increasingly in the widely-used state-of-the-art Weather Research and Forecasting (WRF) model. More recently, the implementation of LCZs into WRF was simplified further, via the improved WUDAPT-to-WRF (W2W) python tool (Demuzere et al., 2022a) and the enhanced WRF model capability since versions 4.3 (Skamarock et al., 2021), enabling the use of the 11 built LCZs without the need for manual code changes.

In order to further facilitate an easier use of LCZs in WRF without relying on any external processing tools, the ingestion of LCZs is now implemented directly into the WRF Preprocessing System (WPS) (Figure 1). This implementation involves two main changes, discussed in more detail below: 1) the development of a hybrid 100-m global land cover product (CGLC-MODIS-LCZ) that uses LCZs to describe the urban areas and 2) the conversion of the CGLC-MODIS-LCZ product to a WPS-compatible data format with an updated WPS table entry for the product.

---

[1] https://www2.mmm.ucar.edu/wrf/users/download/get_sources_wps_geog.html
[2] https://github.com/wrf-model/WPS/blob/develop/geogrid/GEOGRID.TBL.ARW_LCZ

# 2. Data development and implementation

## 2.1 Development of a hybrid 100-m global land cover product (CGLC-MODIS-LCZ)

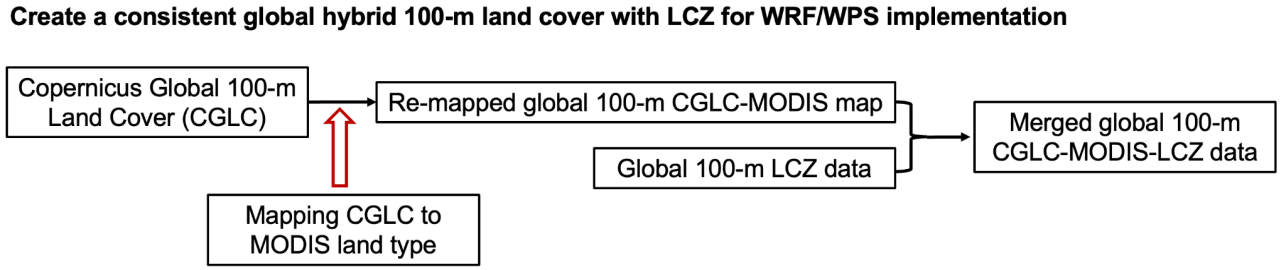


**Figure 1**. Workflow to develop a hybrid 100-m global land cover product (CGLC-MODIS-LCZ).

The global map of LCZs (Demuzere et al., 2022b) is provided on a 100-m spatial resolution that does not match the native resolution of the MODIS-IGBP land cover available by default in WRF (1-km). It is therefore decided to merge the global LCZ map with the 100-m Copernicus Global Land Service Land Cover product (CGLC) representative for the year 2018 (Buchhorn et al., 2021), consistent with the global LCZ map.

First, the CGLC categories are converted to MODIS-IGBP categories using the conversion rules presented in Table 1 based on the biophysical meaning/features of each category. Most categories can be relabeled directly, except the CGLC categories for shrubs (category = 20) and herbaceous vegetation (category = 30), that are broader than the corresponding open/closed shrubland, savanna and grassland categories available in MODIS-IGBP.

**Table 1**: Conversion rules between the CGLC and MODIS-IGBP land cover categories.

| **CGLC** | | **MODIS-IGBP** | |
|---|---|---|---|
| **Value** | **Description** | **Value** | **Description** |
| 0 | Unknown. No or not enough satellite data available. | *22* | *Unclassified** |
| 20 | Shrubs. Woody perennial plants with persistent and woody stems and without any defined main stem being less than 5 m tall. The shrub foliage can be either evergreen or deciduous. | 6, 7, 9 or 10 | See Table 2 for more information<br><br>6: Closed Shrublands<br>7: Open Shrublands: dominated by woody perennials (1-2m height) 10-60% cover.<br>9: Savannas: tree cover 10-30% (canopy >2m).<br>10: Grasslands |
| 30 | Herbaceous vegetation. Plants without persistent stems or shoots above ground and lacking definite firm structure. Tree and shrub cover is less than 10 %. | 6, 7, 9 or 10 | |
| 40 | Cultivated and managed vegetation / agriculture. Lands covered with temporary crops followed by harvest and a bare soil period (e.g., single and multiple cropping systems). Note that perennial woody crops will be classified as the appropriate forest or shrub land cover type. | 12 | Croplands |
| 50 | Urban / built up. Land covered by buildings and other man-made structures. | 13 | Urban and Built-up Land |

| | | | |
|---|---|---|---|
| 60 | Bare / sparse vegetation. Lands with exposed soil, sand, or rocks and never has more than 10 % vegetated cover during any time of the year. | 16 | Barren or Sparsely Vegetated |
| 70 | Snow and ice. Lands under snow or ice cover throughout the year. | 15 | Snow and Ice |
| 80 | Permanent water bodies. Lakes, reservoirs, and rivers. Can be either fresh or salt-water bodies. | 21 | Lake |
| 90 | Herbaceous wetland. Lands with a permanent mixture of water and herbaceous or woody vegetation. The vegetation can be present in either salt, brackish, or freshwater. | 11 | Permanent Wetlands: permanently inundated lands with 30-60% water cover and >10% vegetated cover. |
| 100 | Moss and lichen. | 19 | Mixed Tundra |
| 111 | Closed forest, evergreen needle leaf. Tree canopy >70 %, almost all needle leaf trees remain green all year. Canopy is never without green foliage. | 1 | Evergreen Needleleaf Forest |
| 112 | Closed forest, evergreen broad leaf. Tree canopy >70 %, almost all broadleaf trees remain green year round. Canopy is never without green foliage. | 2 | Evergreen Broadleaf Forest |
| 113 | Closed forest, deciduous needle leaf. Tree canopy >70 %, consists of seasonal needle leaf tree communities with an annual cycle of leaf-on and leaf-off periods. | 3 | Deciduous Needleleaf Forest |
| 114 | Closed forest, deciduous broadleaf. Tree canopy >70 %, consists of seasonal broadleaf tree communities with an annual cycle of leaf-on and leaf-off periods. | 4 | Deciduous Broadleaf Forest |
| 115 | Closed forest, mixed. | 5 | Mixed Forests |
| 116 | Closed forest, not matching any of the other definitions. | 5 | Mixed Forests |
| 121 | Open forest, evergreen needle leaf. Top layer-trees 15-70 % and second layer- mixed of shrubs and grassland, almost all needle leaf trees remain green all year. Canopy is never without green foliage. | 8 | Woody Savannas: tree cover 30-60% (canopy >2m). |
| 122 | Open forest, evergreen broad leaf. Top layer-trees 15-70 % and second layer- mixed of shrubs and grassland, almost all broadleaf trees remain green year round. Canopy is never without green foliage. | 8 | Woody Savannas: tree cover 30-60% (canopy >2m). |
| 123 | Open forest, deciduous needle leaf. Top layer-trees 15-70 % and second layer- mixed of shrubs and grassland, consists of seasonal needle leaf tree communities with an annual cycle of leaf-on and leaf-off periods. | 8 | Woody Savannas: tree cover 30-60% (canopy >2m). |

| | | | |
|---|---|---|---|
| 124 | Open forest, deciduous broadleaf. Top layer- trees 15-70 % and second layer- mixed of shrubs and grassland, consists of seasonal broadleaf tree communities with an annual cycle of leaf-on and leaf-off periods. | 8 | Woody Savannas: tree cover 30-60% (canopy >2m). |
| 125 | Open forest, mixed. | 8 | Woody Savannas: tree cover 30-60% (canopy >2m). |
| 126 | Open forest, not matching any of the other definitions. | 8 | Woody Savannas: tree cover 30-60% (canopy >2m). |
| 200 | Oceans, seas. Can be either fresh or salt-water bodies. | 17 | Water (like oceans) |

** Unclassified pixels are replaced by the modal class of the neighboring pixels. Neighbors are initially taken from a 3x3 window centered on the pixel of interest, a window that is allowed to grow until a modal class is found.*

For the CGLC categories 20 and 30, additional rules are applied, based on the MODIS-IGBP category descriptions that are applied on the land cover fractions available within the CGLC product (Masiliūnas et al., 2021) (Table 2).

**Table 2**: Additional conversion rules for the CGLC categories 20 and 30.

| **MODIS-IGBP** | | **Rules based on CGLC land cover fraction** |
|---|---|---|
| **Value** | **Description** | **Description** |
| 10 | Grasslands | ● Grass fraction > 50%<br>● Tree/Shrub fraction ≤ 10% |
| 6 | Closed Shrublands | ● Shrub + Tree fractions > 60%<br>● Shrub > 30%<br>● Not MODIS-IGBP 10 |
| 9 | Savannas: tree cover 10-30% (canopy >2m). | ● 10% ≤ Tree fraction < 30%<br>● Not MODIS-IGBP 10 and 6 |
| 7 | Open Shrublands: dominated by woody perennials (1-2m height) 10-60% cover. | ● All remaining pixels |

Overall, the above two steps generate the converted global 100-m CGLC-MODIS product. The urban / built up CGLC-MODIS pixels are further replaced by the corresponding built-up LCZ labels (Table 3). The pixels that are identified as built-up in LCZ yet as natural in CGLC keep their natural CGLC-MODIS category value. In the rare cases that CGLC-MODIS pixels are identified as urban yet as natural land cover classes in LCZ, the natural LCZ class is converted to the corresponding CGLC-MODIS class.

**Table 3**: Introducing LCZ labels as CGLC-MODIS categories

| **CGLC-MODIS** | **LCZ** | **Action** |
|---|---|---|
| Urban | Urban | Urban (built LCZs: 51 - 61, equalling LCZ 1 - 10 and E) |
| Urban | Natural | Natural LCZ to CGLC-MODIS class:<br>● LCZ A → CGLC-MODIS 5<br>● LCZ B → CGLC-MODIS 12<br>● LCZ C → CGLC-MODIS 7<br>● LCZ D → CGLC-MODIS 10<br>● LCZ E → CGLC-MODIS 16 |

| | | |
|---|---|---|
| | | • LCZ F → CGLC-MODIS 16<br>• LCZ G → CGLC-MODIS 17 |
| Natural | Urban | CGLC-MODIS Natural |
| Natural | Natural | CGLC-MODIS Natural |

As indicated in Table 1, the CGLC product also contains pixels that are *Unknown* (category 0), as there was no or not enough satellite data available to provide a land cover class (Buchhorn et al., 2021). These pixels are relabeled with the mode from its neighboring pixels, using a 3x3 window centered around the pixels of interest. In case no valid (= non-ocean) modal category was found with this initial window size, the size of the window increased until a valid mode was available.

The general workflow of the aforementioned development processes is demonstrated in Figure 1 above. The final CGLC-MODIS-LCZ product is depicted in Figure 2, and the underlying data is available from Demuzere et al. (2023). Note that this product does not contain the Cropland/Natural Vegetation Mosaic (category = 14), Wooded Tundra (category = 18), and Barren Tundra (category = 20) categories that are available in the default WRF-MODIS data.

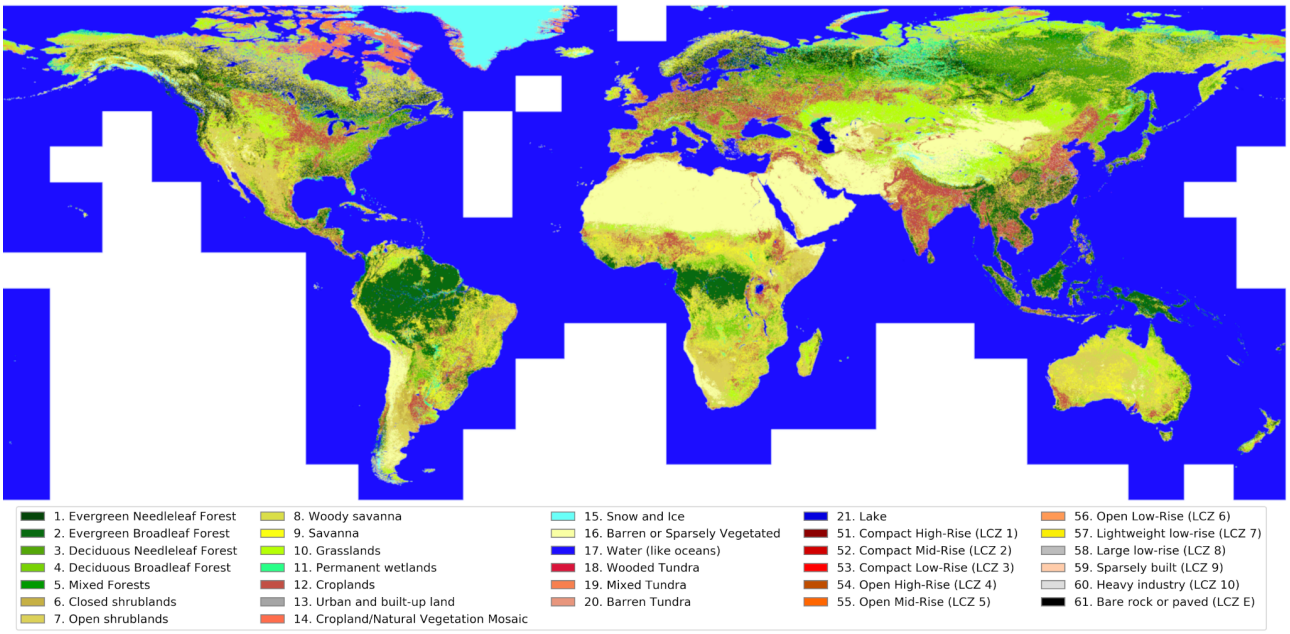


**Figure 2**: The hybrid 100-m global CGLC-MODIS-LCZ land cover map for WRF/WPS. The white areas/blocks are ocean and polar pixels that are not contained in this product. The ocean pixels are excluded to reduce data size and the polar pixels are excluded due to the limited coverage from the original CGLC product. In WRF/WPS, the ocean and polar pixels will be filled with the default MODIS data in WPS automatically if they are included in the users' study domain.

## 2.2 Implementing the CGLC-MODIS-LCZ data into WRF/WPS

To further implement the CGLC-MODIS-LCZ data product developed above into the WRF/WPS system (see Figure 3 for a summary), we conducted the following procedures: (1) converting the raw CGLC-MODIS-LCZ data into tiled binary files required by WPS, (2) creating an index file with the data map information to allow WPS to look for information of a specific study area across all the tiled binary files, (3) adding a new GEOGRID.TBL table entry[2] to define the CGLC-MODIS-LCZ dataset path and averaging rules (nearest neighbor interpolation) with a higher priority over the default WPS MODIS dataset. Thus, if users specify the "cglc_modis_lcz" in the WPS namelist (namelist.wps) for geog_data_res = "cglc_modis_lcz+default", WPS will automatically process the

global 100-m CGLC-MODIS-LCZ data to generate land cover and LCZs over the study domain with relative fractions of each land type for each grid and the dominant land type for each grid. As mentioned above, for some ocean and polar areas that are not covered by CGLC-MODIS-LCZ, they will be filled with the default WPS MODIS land cover map by WPS. The CGLC-MODIS-LCZ data capability is officially released in WRF/WPS version 4.5. The data is archived in the NCAR supercomputer and also available for downloading online[1] (look for "CGLC-MODIS-LCZ_100m").

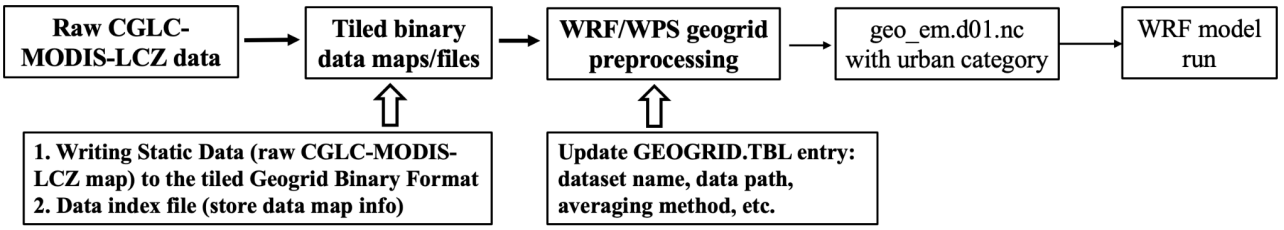


**Figure 3**. Workflow to implement the CGLC-MODIS-LCZ static data into WRF/WPS.

The LCZ modeling capability has been implemented in WRF since version 4.3 to allow the easy use of this CGLC-MODIS-LCZ data. Figure 4 summarizes the LCZ numbers defined in the WRF-urban modeling system, which starts from 51 to 61. Note that before WRF version 4.4.2, the LCZ numbers in WRF are 31 to 41, which however overlaps with the last few land cover classes in the 40-category National Land Cover Dataset (NLCD). Thus, since WRF version 4.4.2, this bug has been fixed to have LCZ numbers from 51~61, corresponding to the original WUDAPT LCZ of 1 to 10 and E (Figure 4).

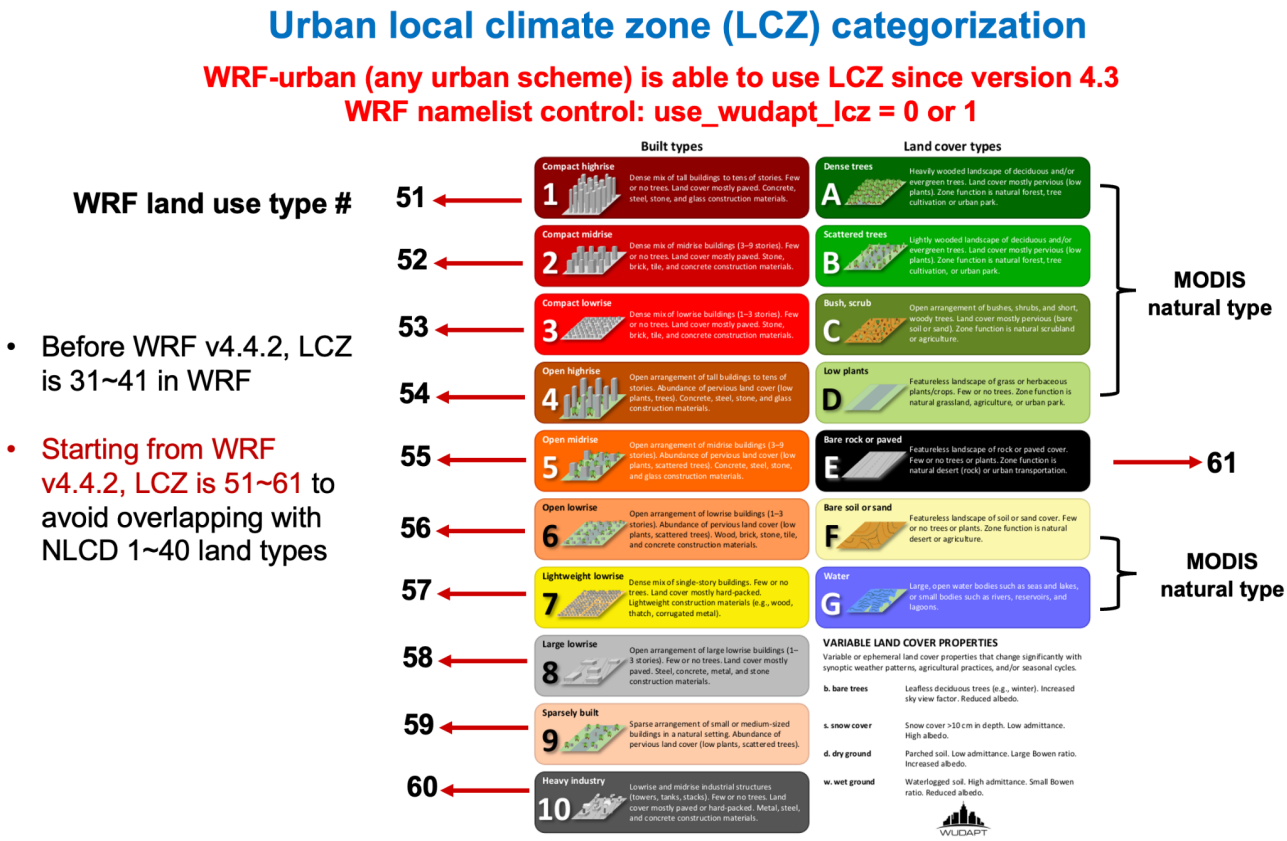


**Figure 4**. Demonstration of LCZ numbering in WRF compared to the original WUDAPT LCZ numbering (Stewart and Oke, 2012; Demuzere et al., 2020).

## 3. Conclusions

- We have developed a hybrid global 100-m CGLC-MODIS-LCZ land cover dataset for WRF based on the the Copernicus Global Land Service Land Cover (CGLC) product resampled to MODIS IGBP classes (CGLC-MODIS), and the global map of Local Climate Zones (LCZ) to describe the urban and built-up land surface.
- The CGLC-MODIS-LCZ dataset has been implemented into the WRF/WPS version 4.5 and is also available for downloading online (https://www2.mmm.ucar.edu/wrf/users/download/get_sources_wps_geog.html)
- Users interested in the original MODIS land cover data with LCZ information can also use the W2W tool provided by Demuzere et al. (2022a).

## References


Brousse O, Martilli A, Foley M, Mills G, Bechtel B. (2016). WUDAPT, an efficient land use producing data tool for mesoscale models? Integration of urban LCZ in WRF over Madrid. Urban Clim. 17(August):116-134. doi:10.1016/j.uclim.2016.04.001.

Buchhorn, M., Smets, B., Bertels, L., De Roo, B., Lesiv, M., Tsendbazar, N.-E., Li, L., Tarko, A. (2020). Copernicus Global Land Service: Land Cover 100m: version 3 Globe 2015-2019: Product User Manual (Dataset v3.0, doc issue 3.4). Product User Manual; Zenodo, Geneve, Switzerland, doi: 10.5281/zenodo.3938963.
Ching J, Mills G, Bechtel B, et al. (2018). WUDAPT: An Urban Weather, Climate, and Environmental Modeling Infrastructure for the Anthropocene. Bull Am Meteorol Soc. 99(9):1907-1924. doi:10.1175/BAMS-D-16-0236.1.

Demuzere M, Hankey S, Mills G, Zhang W, Lu T, Bechtel B. (2020): Combining expert and crowd-sourced training data to map urban form and functions for the continental US. Nat Sci Data 7:1-13. doi:10.1038/s41597-020-00605-z.

Demuzere M, Argüeso D, Zonato A, Kittner J. (2022a) W2W: A Python package that injects WUDAPT's Local Climate Zone information in WRF. J Open Source Softw. 2022a;7(76):4432. doi:10.21105/joss.04432.

Demuzere M, Kittner J, Martilli A, et al. (2022b) A global map of local climate zones to support earth system modelling and urban-scale environmental science. Earth Syst Sci Data. 14(8):3835-3873. doi:10.5194/essd-14-3835-2022.

Demuzere M, He C, Martilli A, Zonato A (2023). A hybrid 100-m global land cover dataset with Local Climate Zones for WRF. [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7670653

Martilli, A., Brousse, O., & Ching, J. (2016). Urbanized WRF modeling using WUDAPT. https://www.wudapt.org/wp-content/uploads/2016/05/Urbanized-WRF-modeling-using-WUDAPT-web-version-March2016.pdf [last accessed 23/02/2023].

Masiliūnas D, Tsendbazar NE, Herold M, Lesiv M, Buchhorn M, Verbesselt J. (2021). Global land characterisation using land cover fractions at 100 m resolution. Remote Sens Environ. 259(April). doi:10.1016/j.rse.2021.112409.

Skamarock WC, Klemp JB, Dudhia JB, et al. A Description of the Advanced Research WRF Model Version 4.3. Boulder, CO; 2021. doi:10.5065/1dfh-6p97.

Stewart ID, Oke TR. (2012). Local Climate Zones for Urban Temperature Studies. Bull Am Meteorol Soc. 93(12):1879-1900. doi:10.1175/BAMS-D-11-00019.1.