# Documentation for the Yale Center for Earth Observation (YCEO) Surface Urban Heat Islands, Version 4, 2003-2018

#### September 2023

Chakraborty, T.,<sup>1</sup> X. Lee <sup>1</sup> <sup>1</sup>Yale School of the Environment, Yale University, New Haven, CT, 06511 USA

#### Abstract

This document outlines the basic methodology and inputs used to construct the Yale Center for Earth Observation (YCEO) Surface Urban Heat Islands, Version 4, 2003-2018, along with previous use cases, limitations, and use constraints.

**Data set citation:** Chakraborty, T., and X. Lee. 2023. Yale Center for Earth Observation (YCEO) Surface Urban Heat Islands, Version 4, 2003-2018. Palisades, New York: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/s5m5-zk14. Accessed DAY MONTH YEAR.

**Suggested citation for this document:** Chakraborty, T., and X. Lee. 2023. Documentation for the Yale Center for Earth Observation (YCEO) Surface Urban Heat Islands, Version 4, 2003-2018. Palisades, New York: NASA Socioeconomic Data and Applications Center (SEDAC). <u>https://doi.org/10.7927/bb9y-9h12</u>. Accessed DAY MONTH YEAR.

We appreciate feedback regarding this data set, such as suggestions, discovery of errors, difficulties in using the data, and format preferences. Please contact:

NASA Socioeconomic Data and Applications Center (SEDAC) Center for International Earth Science Information Network (CIESIN) Columbia University Phone: 1 (845) 365-8920 Email: ciesin.info@ciesin.columbia.edu

### Contents

I.	Introduction	2			
II.	Data and Methodology	2			
III.	Data Set Description(s)				
IV.	How to Use the Data	6			
V.	Potential Use Cases	6			
VI.	Limitations	6			
VII.	Acknowledgments				
VIII.	Disclaimer	7			
IX.	Use Constraints				
X.	Recommended Citation(s)	8			
XI.	Source Code				
XII.	References	8			
XIII.	Documentation Copyright and License 1	1			
Appendix 1. Data Revision History					
Appendix 2. Contributing Authors & Documentation Revision History					

# I. Introduction

The Yale Center for Earth Observation (YCEO) Surface Urban Heat Islands, Version 4, 2003-2018 data set includes annual, summertime, and wintertime Surface Urban Heat Island (SUHI) intensities for daytime and nighttime for over 10,000 global urban extents based on data from Kelso and Patterson (2018). This global SUHI data set was created using the Simplified Urban-Extent (SUE) algorithm and is available at the pixel and urban cluster-levels (i.e. at the level of larger urban agglomerations). The monthly composites are only available as urban cluster means. A summary of older versions, including changes from the data set created and analyzed in the originally published manuscript (Chakraborty and Lee, 2019) can be found on the Yale Center for Earth Observation (YCEO) website.<sup>1</sup> The data set can also be explored using the Global Surface UHI Explorer web application.<sup>2</sup>

The SUE algorithm used to generate the data set is described in Chakraborty and Lee (2019). A short summary is provided here.

# II. Data and Methodology

### Input data

The current version of the data set is created by combining urban extent data from Natural Earth (Kelso and Patterson, 2018), land cover data from the European Space

<sup>&</sup>lt;sup>1</sup> See: <u>https://yceo.yale.edu/research/global-surface-uhi-explorer</u>

<sup>&</sup>lt;sup>2</sup> See: <u>https://yceo.users.earthengine.app/view/uhimap</u>

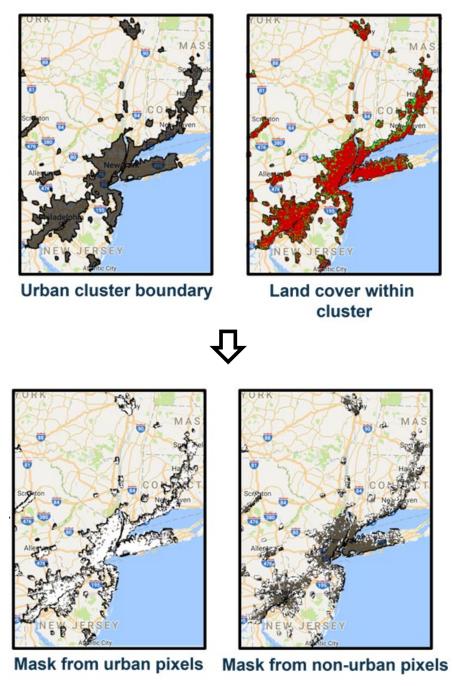
Agency Climate Change Initiative (ESA CCI) (Bontemps et al., 2013), and Land Surface Temperature (LST) estimated from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensors on board NASA's Terra and Aqua satellites (Wan, 2006).

#### Methods

The quantification of the Surface Urban Heat Island (SUHI) intensity requires designation of urban and rural references for comparison. Instead of using urban buffers as the rural reference as done in many previous studies (CIESIN, 2016; Clinton and Gong, 2013; Peng et al., 2011; Yao et al., 2019), land cover data was used to choose rural references from within each urban cluster.<sup>3</sup> Water pixels and pixels with elevation differences greater than 50 meters from the median of the urban pixels in the cluster (as designated by the Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010) (Danielson and Gesch, 2011)) are removed when defining the rural reference for each cluster. Figure 1 shows an example of this delineation of urban and rural pixels for an example urban cluster.

In the original published manuscript (Chakraborty and Lee, 2019) this delineation was done using the MODIS land cover data (MCD12Q1.051) (Cover and Change, 1999), but this version uses the European Space Agency Climate Change Initiative (ESA CCI) land cover data from 2003 to 2018 (Bontemps et al., 2013) (see Appendix). Note that the SUHI calculations are done for urban clusters and their pixels, not for individual cities. This is because there are a host of issues with defining urban areas at broad scales (Chakraborty and Lee, 2019; Potere and Schneider, 2007) and administrative boundaries do not matter for the SUHI (Chakraborty et al., 2020). The cluster-mean SUHI intensity is the difference in mean LST of the urban pixels and the rural pixels within each cluster. The pixel-level SUHI is the difference in LST between the MODIS-derived LST and the mean LST of the non-urban pixels. The 8-day composite LST products (MOD11A2 and MYD11A2 for Terra) are used after only considering pixels with LST error of less than or equal to 3°C. Since Aqua and Terra have different overpass time (~1:30 am and ~1:30 pm for Aqua, ~10:30 am and ~10:30 pm for Terra), the Aqua and Terra products were combined to create daytime and nighttime values.

<sup>&</sup>lt;sup>3</sup> Many UHI data sets compare the temperature within an urban area to the temperatures surrounding the urban area, say with a 5 or 10 kilometer buffer around the urban area. The goal here was to solve the issues with the lack of standardization of buffer width and direction for coastal cities and cities affected by the footprint of the SUHI of a neighboring city. Moreover, not generating buffers for each urban cluster makes the algorithm much easier to implement. The authors compare results for different times against buffer-based methods (see Table 2 of Chakraborty and Lee, 2019) and find that the values from this method are very similar to those from 5 kilometer buffers and lower than 10 kilometer buffers.



### (masks in white)

**Figure 1**: Example of how urban and rural pixels are separated in the SUE algorithm before SUHI calculations for New York City and surrounding urban clusters.

# **III.** Data Set Description(s)

The Yale Center for Earth Observation (YCEO) Surface Urban Heat Islands, Version 4, 2003-2018 data set consists of the following six components:

- 1. UHI\_all\_averaged: Image containing cluster-mean composite daytime and nighttime SUHI intensity averaged across 2003 to 2018 for annual, summer and winter periods (the latter based on hemisphere<sup>4</sup>).
- 2. UHI\_monthly\_averaged: Image containing cluster-mean monthly composites of daytime and nighttime SUHI intensity.
- 3. UHI\_yearly\_averaged: Image collection of cluster-mean yearly composites of daytime and nighttime SUHI intensity from 2003 to 2018.
- 4. **UHI\_yearly\_pixel:** Image collection of spatially disaggregated (nominal scale of 300 meters) annual daytime and nighttime SUHI intensity from 2003 to 2018.
- 5. **Summer\_UHI\_yearly\_pixel:** Image collection of spatially disaggregated (nominal scale of 300 meters) summertime daytime and nighttime SUHI intensity from 2003 to 2018.
- 6. Winter\_UHI\_yearly\_pixel: Image collection of spatially disaggregated (nominal scale of 300 meters) wintertime daytime and nighttime SUHI intensity from 2003 to 2018.

#### Data set web page:

SEDAC URL: <u>https://sedac.ciesin.columbia.edu/data/set/sdei-yceo-sfc-uhi-v4</u> Permanent URL: <u>https://doi.org/10.7927/s5m5-zk14</u> Global Surface UHI Explorer information page: <u>https://yceo.yale.edu/research/global-</u> surface-uhi-explorer

Global Surface UHI web application: https://yceo.users.earthengine.app/view/uhimap

### Data set format:

The data are available in raster and vector formats as a downloadable zip file. The downloadables are compressed zip files containing: 1) GeoTIFFs or Shapefile (SHP), and 2) PDF Documentation.

### Data set downloads:

- sdei-yceo-sfc-uhi-v4-uhi-all-averaged-geotiff
- sdei-yceo-sfc-uhi-v4-uhi-monthly-averaged-geotiff
- sdei-yceo-sfc-uhi-v4-uhi-yearly-averaged-2003-2018-geotiff
- sdei-yceo-sfc-uhi-v4-uhi-yearly-pixel-2003-2018-geotiff
- sdei-yceo-sfc-uhi-v4-summer-uhi-yearly-pixel-2003-2018-geotiff
- sdei-yceo-sfc-uhi-v4-winter-uhi-yearly-pixel-2003-2018-geotiff
- sdei-yceo-sfc-uhi-v4-urban-cluster-means-shp

<sup>&</sup>lt;sup>4</sup> June, July, and August are the months considered for Northern Hemisphere summer and December, January, and February for <u>Southern Hemisphere</u> summer. It is the opposite for winter.

### IV. How to Use the Data

The data are designed for use in a Geographic Information System (GIS) package such as QGIS or Esri ArcGIS.

# V. Potential Use Cases

The data set can be used to derive baseline values of SUHI for urban clusters of interest, including seasonal and long-term trends. Previously, the data set has been used to validate conceptual models (Manoli et al., 2020), and the SUE method and this data set have been used in several other studies (Chakraborty et al., 2019, 2020; Chakraborty et al., 2021; Hsu et al., 2018, 2021; Li, Stringer, Chapman, et al., 2021; Li, Stringer, and Dallimer, 2021; Niu et al., 2020; Sun et al., 2020). The data may be a useful starting point for heat-health related studies, though caution is urged due to the difference between SUHI and air temperature or Canopy Urban Heat Island (CUHI) (Chakraborty et al., 2017; Venter et al., 2021) and the importance of humidity on the physiological response to heat (Anderson et al., 2013).

# VI. Limitations

Note that the SUE algorithm defines the SUHI of an urban cluster as the difference in average LST of all urban and all non-urban, non-water pixels within the cluster. As such, the SUHI is a function of both the urban form and the rural environment. Also note that the urban extent data set used as the units of calculation of the SUHI may be outdated or erroneous, especially for less urbanized areas of the world. There are other important uncertainties that should be kept in mind relevant to SUHI and urban Land Surface Temperature (LST):

- 1. SUHI or urban LST can have a large bias at the individual pixel level. This is partly due to urban thermal anisotropy of MODIS-derived LST (Du et al., 2023). As such, aggregating the SUHI to cluster and even regional scale before analysis is suggested.
- 2. Though SUHI is relevant for understanding modification of the surface energy budget due to urbanization, the magnitude of the SUHI can be largely decoupled from air temperature or canopy UHI during daytime (Venter et al., 2021). This decoupling can be even stronger for daytime moist heat stress across scales (Chakraborty et al., 2022). Directly using SUHI as a proxy for the public health consequences of urbanization is not recommended.

Users should be aware of missing data in the data set. There are two kinds of missing data, either at the cluster-scale or at the pixel scale:

- Cluster-scale missing data are due to the absence of urban and rural pixels (as detected by the ESA CCI land cover product) in a cluster, the lack of any usable pixel in a cluster, or the limitation of the urban extent data used in the algorithm.
- Pixel-scale missing data is due to the limitations of the ESA CCI land cover product (the land cover data may not be valid for every individual city on Earth) or the uncertainty of the MODIS LST data (pixels with > 3°C uncertainty are removed).

The urban extent data set does not change over the years, and thus, cannot account for any change in the footprint of the UHI over the years.

# VII. Acknowledgments

The authors acknowledge the Yale Institute for Biospheric Studies (YIBS) for financial support and thank Professor Dana Tomlin for his suggestions on the Google Earth Engine script.

Funding for the dissemination of this data set was provided under the U.S. National Aeronautics and Space Administration (NASA) contract 80GSFC18C0111 for the continued operation of the Socioeconomic Data and Applications Center (SEDAC), which is operated by the Center for International Earth Science Information Network (CIESIN) of Columbia University.

## VIII. Disclaimer

CIESIN follows procedures designed to ensure that data disseminated by CIESIN are of reasonable quality. If, despite these procedures, users encounter apparent errors or misstatements in the data, they should contact SEDAC User Services at <u>ciesin.info@ciesin.columbia.edu</u>. Neither CIESIN nor NASA verifies or guarantees the accuracy, reliability, or completeness of any data provided. CIESIN provides this data without warranty of any kind whatsoever, either expressed or implied. CIESIN shall not be liable for incidental, consequential, or special damages arising out of the use of any data provided by CIESIN.

## IX. Use Constraints

This work is licensed under the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0).

Users are free to use, copy, distribute, transmit, and adapt the work for commercial and non-commercial purposes, without restriction, as long as clear attribution of the source is

provided.

# X. Recommended Citation(s)

Data set(s):

Chakraborty, T., and X. Lee. 2023. Yale Center for Earth Observation (YCEO) Surface Urban Heat Islands, Version 4, 2003-2018. Palisades, New York: NASA Socioeconomic Data and Applications Center (SEDAC). <u>https://doi.org/10.7927/s5m5-zk14</u>. Accessed DAY MONTH YEAR.

#### Scientific publication:

Chakraborty, T., and X. Lee. 2019. A simplified urban-extent algorithm to characterize surface urban heat islands on a global scale and examine vegetation control on their spatiotemporal variability. *International Journal of Applied Earth Observation and Geoinformation*, 74, 269-280. <u>https://doi.org/10.1016/j.jag.2018.09.015</u>.

## XI. Source Code

No source code is provided.

### XII. References

Anderson, G. B., M. L. Bell, and R. D. Peng. 2013. Methods to Calculate the Heat Index as an Exposure Metric in Environmental Health Research. *Environmental Health Perspectives*, 121(10), 1111-1119. <u>https://doi.org/10.1289/ehp.1206273</u>.

Bontemps, S., P. Defourny, J. Radoux, E. Van Bogaert, C. Lamarche, F. Achard, P. Mayaux, M. Boettcher, C. Brockmann, G. Kirches, M. Zülkhe, V. Kalogirou, F. M. Seifert, and O. Arino. 2013. Consistent Global Land Cover Maps for Climate Modelling Communities: Current Achievements of the ESA's Land Cover CCI. *Proceedings of the ESA Living Planet Symposium*, Edinburgh, September 2013. https://www.researchgate.net/publication/304788651\_CONSISTENT\_GLOBAL\_LAND\_COVER\_MAPS\_FOR\_CLIMATE\_MODELLING\_COMMUNITIES\_CURRENT\_AC\_HIEVEMENTS\_OF\_THE\_ESA'\_LAND\_COVER\_CCI.

Chakraborty, T., Z. S. Venter, Y. Qian, and X. Lee. 2022. Lower urban humidity moderates outdoor heat stress. AGU Advances, 3(5), e2022AV000729. https://doi.org/10.1029/2022AV000729.

Chakraborty, T., X. Lee, S. Ermida, and W. Zhan. 2021. On the land emissivity assumption and Landsat-derived surface urban heat islands: A global analysis. *Remote Sensing of Environment*, 265, 112682. <u>https://doi.org/10.1016/j.rse.2021.112682</u>.

Chakraborty, T., A. Hsu, D. Manya, and G. Sheriff. 2019. Disproportionately higher exposure to urban heat in lower-income neighborhoods: A multi-city perspective. *Environmental Research Letters*, 14(10), 105003. <u>https://doi.org/10.1088/1748-9326/ab3b99</u>.

Chakraborty, T., A. Hsu, D. Manya, D. and G. Sheriff. 2020. A spatially explicit surface urban heat island database for the United States: Characterization, uncertainties, and possible applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 168, 74–88. <u>https://doi.org/10.1016/j.isprsjprs.2020.07.021</u>.

Chakraborty, T., and X. Lee. 2019. A simplified urban-extent algorithm to characterize surface urban heat islands on a global scale and examine vegetation control on their spatiotemporal variability. *International Journal of Applied Earth Observation and Geoinformation*, 74, 269–280. <u>https://doi.org/10.1016/j.jag.2018.09.015</u>.

Chakraborty, T., C. Sarangi, and S. N. Tripathi. 2017. Understanding Diurnality and Inter-Seasonality of a Sub-tropical Urban Heat Island. *Boundary-Layer Meteorology*, 163(2), 287-309. <u>https://doi.org/10.1007/s10546-016-0223-0</u>.

Center for International Earth Science Information Network (CIESIN), Columbia University. 2016. *Global Urban Heat Island (UHI) Data Set, 2013*. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4H70CRF.

Clinton, N., and P. Gong. 2013. MODIS detected surface urban heat islands and sinks: Global locations and controls. *Remote Sensing of Environment*, 134, 294–304. https://doi.org/10.1016/j.rse.2013.03.008.

Cover, M. L., and L. -C Change. 1999. MODIS land cover product algorithm theoretical basis document (ATBD) version 5.0. *MODIS Documentation*, 42–47.

Danielson, J. J., and D. B. Gesch. 2011. *Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010)*. US Department of the Interior, US Geological Survey. https://doi.org/10.3133/ofr20111073.

Du, H., W. Zhan, Z. Liu, E. S. Krayenhoff, T. Chakraborty, L. Zhao, L. Jiang, P. Dong, L. Li, F. Huang, S. Wang, and Y. Xu. 2023. Global mapping of urban thermal anisotropy reveals substantial potential biases for remotely sensed urban climates. Science Bulletin. https://doi.org/10.1016/j.scib.2023.06.032.

Hsu, A., N. Alexandre, J. Brandt, T. Chakraborty, S. Comess, A. Feierman, T. Huang, S. Janaskie, D. Manya, M. Moroney, N. Moyo, R. Rauber, G. Sherriff, R. Thomas, J. Tong, Y. Xie, A. Weinfurter, and Z. Yeo. 2018. *Urban Environment & Social Inclusion Index*. Yale University, New Haven, CT. <u>https://datadrivenlab.org/urban</u>.

Hsu, A., G. Sheriff, T. Chakraborty, and D. Manya. 2021. Disproportionate exposure to urban heat island intensity across major US cities. *Nature Communications*, 12(1), 2721. https://doi.org/10.1038/s41467-021-22799-5.

Kelso, N. V., and T. Patterson. 2018. *Urban areas: Cultural vector (1:10m)*. https://www.naturalearthdata.com/downloads/10m-cultural-vectors/.

Li, X., L. C. Stringer, S. Chapman, and M. Dallimer. 2021. How urbanisation alters the intensity of the urban heat island in a tropical African city. *PLoS One*, 16(7), e0254371. https://doi.org/10.1371/journal.pone.0254371.

Li, X., L. C. Stringer, and M. Dallimer. 2021. The Spatial and Temporal Characteristics of Urban Heat Island Intensity: Implications for East Africa's Urban Development. *Climate*, 9(4), 51. <u>https://doi.org/10.3390/cli9040051</u>.

Manoli, G., S. Fatichi, E. Bou-Zeid, and G. G. Katul. 2020. Seasonal hysteresis of surface urban heat islands. *Proceedings of the National Academy of Sciences*. <u>https://doi.org/10.1073/pnas.1917554117</u>.

Niu, L., R. Tang, Y. Jiang, and X. Zhou. 2020. Spatiotemporal patterns and drivers of the surface urban heat island in 36 major cities in China: A comparison of two different methods for delineating rural areas. *Sustainability*, 12(2), 478. https://doi.org/10.3390/su12020478.

Peng, S., S. Piao, P. Ciais, P. Friedlingstein, C. Ottle, F. -M. Bréon, H. Nan, L. Zhou, and R. B. Myneni. 2011. Surface urban heat island across 419 global big cities. *Environmental Science & Technology*, 46(2), 696–703. https://doi.org/10.1021/es2030438.

Potere, D., and A. Schneider. 2007. A critical look at representations of urban areas in global maps. *GeoJournal*, 69(1–2), 55–80. <u>https://doi.org/10.1007/s10708-007-9102-z</u>.

Sun, Y., S. Wang, and Y. Wang. 2020. Estimating local-scale urban heat island intensity using nighttime light satellite imageries. *Sustainable Cities and Society*, 57, 102125. <u>https://doi.org/10.1016/j.scs.2020.102125</u>.

Venter, Z. S., T. Chakraborty, and X. Lee. 2021. Crowdsourced air temperatures contrast satellite measures of the urban heat island and its mechanisms. *Science Advances*, 7(22): eabb9569. <u>https://doi.org/10.1126/sciadv.abb9569</u>.

Wan, Z. 2006. *MODIS Land Surface Temperature Products Users' Guide*. Institute for Computational Earth System Science, University of California: Santa Barbara, CA, USA, 805. <u>https://lpdaac.usgs.gov/documents/447/MOD11\_User\_Guide\_V4.pdf</u>.

Yao, R., L. Wang, X. Huang, W. Gong, and X. Xia. 2019. Greening in rural areas increases the surface urban heat island intensity. *Geophysical Research Letters*, 46(4), 2204–2212. <u>https://doi.org/10.1029/2018GL081816</u>.

# XIII. Documentation Copyright and License

Copyright © 2023. The Trustees of Columbia University in the City of New York. This document is licensed under a Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/).

# Appendix 1. Data Revision History

The original Version 1 of the data set is the one described in Chakraborty and Lee (2019).

### **Updates in Version 2**

The second version of the data set represents a more realistic implementation of the Simplified Urban-Extent (SUE) algorithm using more recent land use data and leads to a reduction in the mean Urban Heat Island (UHI) intensities (by roughly 20%) compared to the version developed in Chakraborty and Lee, 2018. The methodology has been modified in several ways, namely:

- The European Space Agency Climate Change Initiative (ESA CCI) land cover data are used instead of the MODIS land use/land cover data. In the previous version, urban and rural pixels were chosen on the basis of the MODIS land cover classification (MCD12Q1.051), where urban pixels remain unchanged over the years. In this updated version, urban and rural pixels in each urban cluster are decided on the basis of the ESA land cover classification, which is updated every year and is available at a higher resolution (300 meter vs 500 meter for MODIS).
- 15-year means (for monthly, annual, summer and winter) are calculated based on the UHI estimates for each year taking into account the land cover data for the corresponding year. Previously, the most recent (2013) MODIS land cover data were used for the calculations, since they did not change over time. This leads to negligible changes in the UHI estimates (< 1%).
- Instead of filtering out urban clusters where the mean elevation difference between the urban and non-urban, non-water pixels are greater than 50 meters, the rural reference to calculate the UHI was modified by masking out the pixels that have an elevation difference of more than 50 meters with the median of the elevation of the urban pixels. In addition, clusters are no longer filtered-out based

on the percentage of urban area in a cluster. By preventing the filtering out of entire urban clusters and using higher resolution land cover data, this update significantly increases the sample size (from 7,374 to 10,136).

#### **Updates in Version 3**

This version update pertains primarily to improved functionality of the <u>Global Surface</u> <u>UHI Explorer web application</u>. The app now allows users to download the gridded UHI data for a selected urban cluster. Users can scroll down to the bottom of the right-hand panel to extract the data sets for their selection. The masked pixels have a fill value of -999. As also mentioned in the original paper, the SUE algorithm attempts to create a globally consistent UHI database by generalizing the selection of the rural reference. As such, caution should be taken when analyzing the UHI magnitude of individual clusters, especially at the pixel scale.

#### **Updates in Version 4**

The methodology used to create this version remains the same as the third version, but the land cover data used for the years 2016 and 2017 are the recently released year-specific ESA CCI land cover layers for those years (the layer for 2015 was used for 2015, 2016, and 2017 in version 3). In addition, the calculated UHI intensity data for 2018 based on the ESA CCI land cover data for that year is added. The data from 2003 to 2018 are now used to calculate all the yearly, seasonal, and monthly composites.

### **Appendix 2.** Contributing Authors & Documentation Revision History

Revision Date	ORCID	Contributors	Revisions
September 5, 2023	0000-0003-1338-3525 0000-0003-1350-4446	5	This document is the 1 <sup>st</sup> instance of documentation.