Manuscript essd-2022-92, First Revision
Submitted to: Earth System Science Data

A global map of Local Climate Zones to support earth system modelling and urban scale environmental science

Demuzere, M., Kittner, J., Martilli, A., Mills, G., Moede, C.,
Stewart, I. D., van Vliet, J., and Bechtel, B.

Response to Referee #3

July 14 2022

We would like to thank Dr. Jason Ching for his extensive appreciation of our work and the valuable comments on the manuscript. Point-by-point responses to the minor issues are mentioned below, indicated in blue.

Preface to this review: A decade ago at the Croucher Advanced Study Institute in Hong Kong, this Reviewer and Gerald Mills (a coauthors of this paper) reflected upon a presentation by Iain Stewart (also co-author on Local Climate Zones (LCZ), topic of his PhD research. The LCZ is a universal classification scheme that differentiates urban surfaces into different combinations of building form and function features. Together with associated values of urban canopy parameters (UCPs) the LCZ provided the conceptual framework that inspired the startup of WUDAPT, an urban climate community collaborative project. WUDAPT scope is worldwide, in principle the global LCZ map with its companion UCPs provides model inputs that describe the underlying embedded canopy features of the urban boundary layer for any and all cities in the world. This paper describes efforts leading to the generation of the Global LCZ map, an achievement that culminates an effort a decade in the making and satisfies the major goal and is a key milestone of the WUDAPT. This global LCZ map product support “fit for purpose” applications of environmental models capable of addressing urban induced environmental issues exacerbated by climate changes.

Overview: This paper is a significant contribution; it represents the product at 100-meter resolution designed to provide urban canopy data, making possible a means to generate uniformly consistent urban canopy parameters for all cities in the world as inputs to a wide variety of models such as meteorology (e.g., WRF_ Surface energy budgets (e.g., SUEZ), etc. The implication of this result is a capability for making possible fit for purpose (FFP) modeling for assessment at intra-urban scales impacted by climate changes for any and all cities in the world. The LCZ scheme is the cornerstone of this paper; it defines 10 distinct classes each having unique land cover and physical form and function aspects describing the built environment along with 7 other nonurban land cover types. For each of the built classes, there is a corresponding range of values of urban canopy parameters suitable as modeling inputs for urbanized WRF and other environment modeling systems. The research team has been engaged in developing and establishing methods and techniques at the outset of WUDAPT; their creating of an LCZ generator facilitated upscaling LCZ maps based on sets of Training Areas (TA) representing each LCZ class by urban experts for individual cities upscaled to regional/continental scale maps for different regions of the world. Their R&D trajectory has provided the experiential base and creating approaches and methods that extend TA transferability from individual cities to regional and continental maps. This paper describes in detail the adjustments and modifications to the methodology that generated regional-continental LCZ maps thus enabling the creation of this global LCZ map. This achievement fulfills a major objective of the WUDAPT (www.Wudapt.org) Project.

Key Points: The article provides the reader with (i) a concise discourse describing the approach and methods and inputs to generate this product; and (ii) provide suggestions on its utility to supporting modeling and urban scale environmental science. The context of my review will reflect perspectives based on the WUDAPT initiative, a collaborative venture of the urban community. This article is thoughtful, and well organized, I briefly highlight and summarize key points of each section below.

The Title is accurate and appropriate. The focus of this paper in on the development and implementation details of the methodologies that makes possible the resulting global LCZ map. LCZ is an universal urban typology and classification scheme to representing unique properties of form and function variables in the urban canopy layer of cities along with 7 other classes representing non-urban landscapes. The complementary urban canopy parameters values associated with each LCZ class provide a globally consistent framework for data on the form and function of morphological features in urban canopies. This effort extends LCZ maps progressing from original city specific sets to recent mapping of cities and the surrounding non-urban areas within regional domains to global coverage. The impact of this achievement supports urban scale modeling systems and their applications anywhere in the globe, achieving the primary objective of the Level 0 approach of WUDAPT. Implementation procedures to generating the upscaled Global product required appropriate modification and innovations to previous efforts for the regional scale prototypes.

The Introduction provides a perspective on the value and importance of urban science to addressing global climate change issues. The enabling science behind current models is the physics algorithms for the vertical exchange processes of momentum, energy and moisture pollutant emission influenced by the myriad of urban morphological (UM) features. These exchange processes take place in the so-called urban canopy layer which extends from the surface to the top of UM features. Typically, modeling with urban canopy physics requires special sets urban data but heretofore is only available for limited number of cities; thus a large information gap exists which severely limits environmental models as tools intended and needed to addressing climate induced risks issues at urban scales. The introduction of the LCZ framework and its companion range of values of urban canopy parameters (UCPs) is what the Global LCZ map achieved here and generated at 100m resolution makes possible myriad of practical ”fit for purpose” environmental modeling applications at a reasonably fine scale addressing climate change issues impacting weather, climate, air quality at both inter and intraurban scales for each and all cities worldwide.

Section 2 describes in several subsections, specific details of each of the various methodology and approaches employed to generate this global LCZ map. In general, the approach pertinent to each aspect is described in detail, thoroughly, albeit, many supporting technical details were provided in cited references.

<u>Section 2.1Training Areas (TAs) </u>This subsection describes the methods to generate this Global LCZ map. It is based on incorporating TAs of LCZs from various sources, mainly from (a) archived community generated TAs representing hundreds of different cities around the world and additionally (b) a special set prepared for another hundred or more other Regions of Interest (ROIs) cities around the world and (c) from TA samples generated by a unique LCZ Generator employed earlier for regional mapping projects (reference cited). The archived TAs required a curation effort to rectify the issue of unevenness in the quality and physical size of TAs submitted into the WUDAPT archives. Their approach adopted in the curation processing is logical, and sound towards assuring uniformity and consistency in quality of the TAs. It clearly builds upon insights and experience gained in prior efforts including the HUMINEX project, implementing such approach contributing to the successful generating of regional to continental scale LCZ maps. The

efforts described in this section extends the approach used for the regional maps to assure a uniformity in the quality of this global product.
Section 2.2 describes special treatments to extend the handling of the added and supplementary Earth Observations (EO) needed as inputs for the Global LCZ supervised random forest classifier from the regional mapping stage. Here updates or additions to the original 33 Global Earth Observation were incorporated to the LCZ Generator.

Section 2.3: Classification schemes This section describes what the authors call a Lightweight Global Random Forest model based on various pixel-based mapping methods; however, upscaling such methods to the global product was apparently not straightforward. For this, the default LCZ Generator from earlier studies required significant modifications. This was apparently a huge classification challenge (given >10^6 labelled TA and other inputs); it was facilitated by incorporating this dual sequential pathways approach, another innovative advancement.

The important QA assessment builds upon their recent continental scale LCZ mapping efforts; its procedures are based on (a) five (5) traditional accuracy metrics (Section 2.4.1) and (b) incorporating a novel but indirect thematic benchmarking (Section 2.4.2) involving comparing mapped outcomes of several urban canopy parameters (% built, % impervious surface and sum of built and impervious total plan area, building heights and AH (anthropogenic heating) associated with each LCZ class with other sets of global and open source databases reflecting urban form and functions. The level of comparability provided a relative qualitative assurance measure of the outcomes of UCPs associated with the LCZ maps. Clearly, the success of these relative outcomes varied for the different UCP analyzed (Fig 10). In this regard, future effort associated with other independent means including outcomes of UCPs generated by WUDAPT Level 1 and 2 approaches described in Ching et al, (2019) will be helpful, going forward.

Section 3: Results. The Global LCZ map is shown in Fig 4. While the 7 non-urban LCZ classes are not as discriminating in the number of classes as in other mapping schemes, the major value is that all urban areas herein are discriminated into the 10 universally based LCZ scheme, a product consistent for all urban areas in the world, e.g., Figure 5-7, are examples of zoomed LCZ maps for various cities extracted from this Global LCZ maps. It was noted that these sets of cities display and support earlier observations that that each and all cities has its uniquely characteristic LCZ signature (or fingerprint). This is a feature that was apparent and evident in the UCPs generated for the NUDAPT project (Ching et al, 2010), and for LCZ and UCP maps for individual cities studies and from the recent regional LCZ mappings. From such observations, it is probably reasonable to infer that all cities in the world have unique LCZ fingerprints, and by extension, a commensurate unique set of UCPs. This is an important consideration as it provides the rationale and bases of conducting fit-for-purpose intraurban modeling studies based on the Global LCZ map applicable to each and every city in the world. Support for the contention is expressed in Section 4 indicating how this Global LCZ map can serve earth system modeling and urban scale environmental science, and extend its utility to intraurban scale. However, as noted in the paragraph beginning at line 492, much more need to be done for full effectiveness especially at the intraurban scale in this regard. The UCPs provision in LCZ is currently manifested in lookup tables of ranges of values in the UCPs of each LCZ class. Remedies include path forward innovative cyber-based approaches are currently underway to generate block scale gridded UCPs of both form and function parameters in WUDAPT Level 1 and 2 staged efforts already referred to in the reference list (Ching et al. 2019) to complement the global LCZ mapping efforts.

Section 4: this Section, the article highlights and discuss the attributes and impact of the Global LCZ product. Herein, the significance in terms of objectives, and potential impacts and caveats of this study are explored, Since this global map has just been completed, the discussion refers to the

Global LCZ maps support to wide range of potential modeling applications, some already underway, in concert with the WUDAPT perspecitives.
Conclusion section: The results of this Global LCZ map is an important and significant achievement culminating from the collaborative efforts of many activities and voluntary contributions from the urban climate and multidisciplinary science communities evolving and improving after over nearly a decade of efforts. The advances and contribution by this team has been impressive moving the LCZ framework from prototypic city specific mapping to the creation of regional maps and culminating in this impressive global product. Given the widespread and rapidly growing literature on LCZs we can anticipate much interest in this product. For a whole host of reasons, including projections of climate change impacts to enhancing weather extremes, to urbanization dynamics from increased population, the LCZ paradigm, and the various levels of coverage, and certainly, with this Global product provides an important and significant approach towards supporting science-based tools for myriad and wide ranges of modeling application and studies supporting local to international policies that addresses climate change issues.

Data availability: In the last section, this map is available via link provided. While it was not mentioned, I would recommend the authors consider making the map and accessibility to the pixel generated LCZ in the WUDAPT Portal
Please see the response below.

Suggestions on specific points: The following are a list of relatively minor issues and concern indicated below.

Line 120-121.This sentence bears a burden of explanation; It will be necessary to provide objective measures and criteria to establish the objective quality indicators of the RUB produced TA vs other ARC sets.
More details on the filtering procedures for the different sets of TAs are provided in lines 126 to 136. In general, it is important to realise that all TA polygons are used, from the three available sources: RUB, ARC and GEN. Irrespective of the source of the samples, they are subjected to general filters, eg. retaining only the submissions with OA > 50%, or removing too small or too complex TA polygons.
Yet only in one filter a prioritisation is applied. That is when submissions from multiple sources cover the same city. In this case we prioritise RUB > ARC > GEN (as indicated on Line 129). This priority is chosen because 1) it follows the order of the average OA across these sources (RUB: 73%, ARC: 70% and GEN: 65% - after filtering for those with OA > 50%), 2) RUB samples are given priority as these are produced in a more controlled setting, and 3) also ARC samples are a product of a more controlled setting, as most of these TA sets are the outcome of a peer-review paper and/or quality-controlled by WUDAPT-experts as part of the manual quality procedure in the original WUDAPT portal.
Note that the controlled setting for the RUB samples is further explained in the answer to question 1 of reviewer 1: “ Please define what a well-trained student is”.

Line 123, page 6.Revise, eliminate term “old“ in old WUDAPT Portal. This eliminates the need to explain or differentiate the progression of Portal versions.
We appreciate this feedback. We changed “old” to “original” as we do want to make a distinction between the first portal (that is currently no longer maintained nor updated, as indicated on https://www.wudapt.org/the-wudapt-portal/), and its successor, the LCZ Generator.

Line 131, “Third” used here is really “Fourth” as “Third is already used in line 129
The numbering of these sentences has been adjusted.

Line 189:To better understand and appreciate the value of the Lightweight global random forest model, herein, as regards the probability layer, please discuss the takeaway points of the meaning of a high vs low probability.

The lightweight Random Forest model was introduced for computing efficiency since the entire dataset was too large to train one classifier (even with the tremendous computing resources offered by Google Earth Engine). Thus the ensemble-based decision making principle of Random Forest was adapted using an outer split to reduce the problem to several smaller problems, i.e. 50 classification models based on a subset of the data. The effect on the accuracy is unknown but considered to be minor, since the procedure is similar to what RF does internally anyway (with different parameters, sensitivity which was tested in pathway 1), just with our approach only 15 % of the memory is needed for each run. As added value the results from the 50 runs can be interpreted as a classification probability (see also rebuttal Review #1). A high classification probability means that most of the 50 RF models agree (i.e. class is chosen independent of training subset). A low classification probability means that the RF models largely disagree (i.e. results strongly depend on the training subset).

The role of the classification probability layer is described in Section 3.2, Figure 9, and a new Figure H1. In short, this section describes that "*all classification probabilities per LCZ class are in line with the global values, demonstrating the universality of the LCZ typology and the robustness of the classifiers and input features across the urban ecoregions*". It also indicates areas where more work is needed, such as the mapping of LCZ class 7, which is indicated as one of the caveats in the Discussion section.

Line 246 Functional Urban Area (FUA) are indicated by a reference (Schiavina et al. 2019); given the important role of FUAs in discussions in the remaining text it would be highly useful to introduce the key aspects of this FUA framework as in the Section 2.3 and 3.2.
Please note that in-depth information on FUAs is already provided in Appendix C, which will be part of the main manuscript and will thus not be provided as supplementary information:
Functional urban areas (FUAs), as defined by the Organisation for Economic Co-operation and Development (OECD) and the European Union, are sets of contiguous local (administrative) units composed of a city and its surrounding, less densely populated local units that are part of the city's labour market (commuting zone). As such, these units not only offer the opportunity to evaluate more densely built city centres, yet also their sparsely built or natural neighbouring landscapes.

This Appendix is explicitly referenced upon the first usage of FUAs. In order not to clutter the main manuscript with too many details, we have decided to keep this information in this Appendix.

Line 433 pg 24. Clarification of introducing the Term "scaling laws"
This paragraph has been adjusted to better explain the idea of "scaling".

Line 503: Consider elaborating on introducing the term "subclass scheme"
Some information is added to clarify this, yet more importantly the paper from Stewart and Oke (2012) is explicitly mentioned, since it provides much more details in the LCZ subclasses.

Explain "Morphological Gaussian Filter"
In general, many parts of this work built upon previous results, that are described in previously published works. We believe it is not required to repeatedly explain underlying procedures, as long as the original papers describing those are properly referenced. In the case of the morphological gaussian filter, this is also the case, rereferring to Demuzere et al. (2020a).

Highly recommend the Global LCZ Map and data results be incorporated into the WUDAPT Portal.

Since the original WUDAPT portal is no longer maintained nor updated, we do not plan to put the Global LCZ map data there. We do however foresee a number of outlets that will allow interested users to access the data:

1. The official Zenodo data archive: https://zenodo.org/record/6364594. Even though this manuscript is not officially accepted yet, the data has been seen and downloaded 600+ times already [last accessed July 11 2022].
2. Interactive viewer on the LCZ Generator: we are currently designing another tab on the LCZ Generator page that will display the global map in an interactive manner.
3. GIS and other tools: Since we had to develop tiles for 2., these will also be accessible by other software such as QGIS and similar, that support xyz tile servers.
4. Earth Engine: once the manuscript is accepted, the data will be released in the official Earth Engine data repository, as described here: https://developers.google.com/earth-engine/help_dataset_description.

Page 51, Last line of caption to Figure F1; add "detail in" the Table F1.
Good catch! "detail' is added.

Summary: The effort described in this article to creating this Global LCZ map is impressive. This product represents an achievement, culminating a decade of activity and efforts by the urban community through WUDAPT towards acquiring urban canopy layer data for models that provide a means for addressing climate change and urbanization issues for local to regional to global scales. Its paradigm incorporates the LCZ typology thus providing a unified and consistent basis for generating intraurban form and function type data paradigm (WUDAPT Level 0) for the urban canopy layer. This Global LCZ map supports the observation from earlier city specific studies and regional maps, that each city LCZ map signature (e.g. fingerprint) is unique. This "feature" provides a rationale for supporting a wide range of earth systems modeling, applications, and urban scale environmental science, e.g., urban modeling applications in which intraurban scale weather forecasting and assessments can be made based on implementing urban boundary layer parameterizations in models with universally consistent intraurban urban canopy descriptions unique to each and every urban area in the world. The rationale, requisite technical issues to the innovative approach toward generating this global map, the approach and results were well articulated and fully documented. The point to future work advancements some alluded to in the caveats provided of the current LCZ paradigm for balance. In particular, its incorporation into WRF for example through an improved WUDAPT to WRF link is underway as well as efforts along the lines of WUDAPT level 1 and 2 (Ching et al. (2019) will be to provide a pathway towards introducing refined city specific block scale gridded UCPs in future updates.
Thank you Dr. Ching for the supportive words!