Detailed point by point response
to *#Anonymous reviewer 1*

Below are summarize each comment of reviewer 1: first the discussion between reviewer and authors concerning a specific point; second the changes made in the manuscript concerning this point (in this section, all lines refer to the marked-up manuscript version).

Discussion 1

Reviewer (R): The manuscript is written quite well. The model is well decribed, the concept of the manuscript is well organized and it is easy for reading and understanding. It is noticable that authors have experience in modeling as well as in writing the scientific papers. Also GeoClimate software is useful in climate research, and it is open-source software, open of new collaborators and useful for datasets analysis. Also, here is presented that is good tool for the LCZ clasification. Therefore, this manuscript can be considered for publication in this journal.

Also, there are some disadvantages that should be discuss in the future. After the reading this manuscript, the impression remains that it is still sufficient to use the LCZ Generator (WUDAPT), and that GeoClimate does not provide any novelty in defining the LCZ. Maybe for authors it was not the main goal, but maybe readers will expect to see new tool that should be better than old ones. As it was highlighted in the Conclusions...the integration of GeoClimate and WUDAPT tools could make significant improvements in furhter LCZ classification and this should be a next step of the authors.

Authors (A): Thank you to anonymous referee #1 for the comments.

Concerning the discussion about WUDAPT and GeoClimate comparison, as the referee assumed, it is not the purpose of this manuscript. The GeoClimate

LCZ algorithm has been available for the community since several years so it was needed to:

- describe clearly what was the methodology used to determine the LCZ of a given area
- evaluate what would be the difference in using the worldwide available OSM dataset instead of the French BDTopo one.

Some preliminary works (such as in Blond et al. (2023)) have been achieved to compare GeoClimate to WUDAPT on some French cities but they have not been published yet. The main rough observation is that GeoClimate seems more appropriate than WUDAPT for urban areas but less appropriate for rural areas.

Blond, N., Breton, F., Micolier, A., and Mendez, M.: A modeling approach to address building energy consumption and thermal comfort under urban climate change , EGU General Assembly 2023, Vienna, Austria, 24–28 Apr 2023, EGU23-14169, https://doi.org/10.5194/egusphere-egu23-14169, 2023.

R: Dear authors,

OK. For me, your replies are acceptable. I will not have any comments.

- Changes in the manuscript : The last paragraph of the introduction (l. 92 to l. 96) has been modified in order to better clarify the purpose of the comparison between using OSM and BDTopo datasets.

Detailed point by point response to *Jan Geletic*

Below are summarize each comment of Jan Geletic (reviewer 2): first the discussion between reviewer and authors concerning three specific points; second the changes made in the manuscript concerning these points (in this section, all lines refer to the marked-up manuscript version).

Discussion 1

Reviewer (R): You mentioned the widely known LCZ generator, WUDAPT. There is no comparison and I fully accept it (as you stated in reply). But which benefits can we could expect if your algorithm will be used? Is it better, faster, integrated... This study is focused on French cities only, so I am not sure.

Authors (A): For many European and North America cities, we expect a better identification of urban LCZ types using GeoClimate and OSM than using LCZ generator since the degree of completeness of OSM is rather good for building geometry and building types in these regions and also OK concerning building height (see Bernard et al., 2022). However, LCZ generator is probably more appropriate for rural areas since many small tree vegetation patches or types might be missing in OSM. Moreover, there is still a big lack of OSM data such as in many African, Asian or South American countries. As described in Bocher et al. (2021), GeoClimate algorithms can roughly be distinguished in two steps: first data are imported into a well defined and generic data model; second all indicators and classifications are performed using the data previously included in the generic model. Thus GeoClimate might still be used if local vector dataset exist. However, the first step will still be needed to be performed.

R: Thank you for your reaction and references. [This point is] fine for me.

- Changes in the manuscript : A sentence has been added at the end of the first paragraph of section 2.1 (l. 100 to l. 102) to clarify that any data source can be used as input of the GeoClimate LCZ processing chain.

Discussion 2

R: There were presented and discussed differences between BDT and OSM. But which layer is more precise? Did you somehow validate results with reality (e.g., using manually defined samples etc.)? How accurate algorithm is?

A: As discussed in the manuscript section 2.5, the main expectations we may have about the differences between OSM and BDT data are that:

- BDT building height is more accurate than the OSM one since for most of the OSM buildings, this information is simply estimated using a random forest model using BDT as real value for the training (cf. Bernard and al., 2022).
- OSM data has a generally higher land coverage which is mainly due to a better representation of impervious area and vegetation within cities and also a higher building coverage. This information has been verified in our study and is discussed section 3.3. However, the statistics given have been inverted and will be corrected in a next version (37% in BDT against 55% in OSM).

R: Further explanation for point 2 is still needed. As you stated, your paper was submitted as 'Model description paper'. See, please, its detailed definition on GMD website, specifically point below:

Examples of model output should be provided, with evaluation against standard benchmarks, observations, and/or other model output included as appropriate. In this respect, authors are expected to distinguish between

verification (checking that the chosen equations are solved correctly) and evaluation (assessing whether the model is a good representation of the real system). Sufficient verification and evaluation must be included to show that the model is fit for purpose and works as expected. Where evaluation is very extensive, a separate paper focussed solely on this aspect may be submitted.

Can you, please, select one of the cities you classified and compare it with an expert-based classification? Or with a WUDAPT method? This information is important for a potential user; without this information you cannot state that model provides relevant or sufficient results. Attached is a manually classified sample for Brno in the Czech Republic, if you have no own sample.

A: Thank you for your quick reaction.
You are waiting for further explanation concerning the lack of evaluation of the method. We understand and agree that comparing the GeoClimate results to other state-of-the art methods is necessary and as previously said as answer to your “major comment 1”, we will make this comparison in an early coming future.

You state that as a “Model description paper”, our manuscript must contain an evaluation section. You refer to a paragraph where the word “should” is used. Thus this paragraph is not necessarily applicable (below is the distinction made by GMD between "should", "must" and "may", this paragraph being located at the begining of the “manuscript type” page of GMD).

'In the following, "must" means that the stated actions are required, and the paper cannot be published without them; "should" means that we encourage the action, but papers can still be published if the criteria are not met; "may" means that the action may be carried out by the authors or reviewers, if they so wish.'

We have made our best to fullfill all the “must”, most of the “should” and some of the “may” but we have not fullfilled the evaluation part which is a “should”. Concerning this one, we need to reaffirm our position. There are two objectives within this manuscript: the first (the main one) aims at describing accurately what is performed within the GeoClimate LCZ algorithm; the second is to illustrate the differences obtained using two datasets that are currently automatically usable with GeoClimate. Still, the

"evaluation" (which would be more a comparison in the case of LCZ) of GeoClimate using the LCZ map produced by an other method on a single territory could be performed and added to the manuscript as you proposed. However, this involves:

- A third objective to our article (or at least a new section) which might make the manuscript hard to follow and too long
- Having the LCZ map of a French city since the article focus on French cases
- The results of the "evaluation" (comparison) would be valid only for this single territory. Thus it will only be a sort of illustration more than an evaluation or an interesting comparison where general conclusions can be made.

In order to have a proper evaluation, a solution would be to compare the GeoClimate method to an other one which has been applied to many different French locations. However, this raises two issues:

- The manuscript would get really long
- For now and at our knowledge, only the WUDAPT method has been applied at such scale and as previously discussed in our answer to your "major comment 1", we have planned to do this comparison in a separate article

Our point of view is that this manuscript is a good base for future comparisons involving the GeoClimate algorithm. This first article describes the algorithm and the limitations of using one of the two datasets (BDT or OSM). As a consequence, for French applications, the current optimized dataset would be to use a combination of OSM and BDT. Moreover, we have now evaluated the limitation of using the OSM data (which miss the building height) thanks to a reference dataset (BDT where building height accuracy is known). This preliminary knowledge was at our point of view necessary before to evaluate further the GeoClimate method using OSM on other territories than France.

- Changes in the manuscript : The last paragraph of the introduction (l. 92 to l. 96) has been modified in order to better clarify the purpose of the comparison between using OSM and BDTopo datasets.

Discussion 3

R: Geodatabase in OSMs is based on community - quality is definitely not the same worldwide. Moreover, there is much more data available for OSM editors. I am quite skeptical of the trustworthiness of OSM in some parts of the world. Despite that fact, in Europe exists freely-available sources with a strong potential for delineation methods - Copernicus Land Monitoring Services. Do you know about these datasets with a spatial resolution in tens of meters (e.g., Urban Atlas, European Settlement Map, Imperviousness etc.)? It would be interesting novelty…

A: As previously discussed, there is indeed almost no OSM data in some part of the world and thus OSM is not applicable there yet. However, as previously explained also, GeoClimate is a two steps approach where the first step concerns data import into the generic GeoClimate data model. Thus some other datasets can be merged during this first step (even though it might mean a considerable amount of work depending on data completeness) and then be used in the presented GeoClimate algorithm (second step of the GeoClimate processing chain). But the first objective of this article is to share the methodology used by GeoClimate to identify the LCZ (so the second GeoClimate step), not to show combination of data that can be used in GeoClimate (first GeoClimate step). However we have illustrated the behavior of the algorithm depending on two different datasets. In the future, a potential work could indeed be to try to merge different datasets to obtain the best land coverage and thus LCZ classification but this has to be done in a future work, this one being the description of the LCZ methodology used by GeoClimate.

R: Thank you for your reaction and references. [This point is] fine for me.

- Changes in the manuscript : Slight modifications have been performed at the beginning of section 2.1 (l. 103 to l. 113) to better highlight the fact that any dataset can be used as input.