Earth Syst. Sci. Data Discuss., referee comment RC2
https://doi.org/10.5194/essd-2022-92-RC2, 2022

Comment on essd-2022-92

Anonymous Referee #2

Referee comment on "A global map of local climate zones to support earth system modelling and urban-scale environmental science" by Matthias Demuzere et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2022-92-RC2, 2022

General comments

This work describes a new dataset of land cover types (10 urban – 7 natural) using Local Climate Zones at the global scale. Work is clearly described, evaluated, and presented. The associated dataset is of high quality, and I expect will become a landmark data source for the community.

The discussion on “accuracy” vs “robustness” could be improved (see specific comments). Additionally, there is no acknowledgment that LCZ training polygons are susceptible to human errors (again see specific comments).

Section 3.2 and Figure 10 show that the correlation R2 for building height is only ~0.5, however this is only very briefly mentioned in results, and not mentioned elsewhere (e.g. discussion/conclusion/abstract). So, while 2D information like lambda_B appears to be very well captured, 3D information remains a significant limitation. This is a key result and its implications should be discussed more thoroughly.

A lower reliance on acronyms would assist the casual reader. For example Figures 9 and 10 are not decipherable without referring to other sections of the text.

However, overall, an impressive body of work.

Specific comments

Line 41: “Earth System Models (ESMs) have only recently evolved to accommodate urban-scale landscapes, even though the parameters that are used by ESMs to these landscapes are limited in scope”

Some global climate models have had integrated urban canyon models for over a decade (e.g. CLMU in CESM). I’m not sure if these are ESMs (ESM relates to the carbon cycle, not the global scale, some readers may misinterpret this). I think safer/clearer to say many global-scale models ignore urban landscapes or represent them simply.

Line 120: suggest removing “well-trained” as subjective.

Line 127: “only the best submission is retained” what distinguishes a “best” submission?

Line 128: How is accuracy determined?

Section 2.4.1: I would describe this as a test of robustness, not accuracy, as this does not test whether the classifications are correct, just whether they change with different inputs. This method also assumes that training areas are accurate, but TAs are classified subjectively by humans. True accuracy can be tested with building resolving spatial datasets. However, I accept this “accuracy” terminology has been established elsewhere in the literature, but a comment to clarify accuracy vs robustness would assist readers.

Line 200: “The overall accuracy denotes the percentage of correctly classified pixels.” As described above, the method does not assess whether pixels are classified correctly, only how often they are unchanged (and potentially remain incorrect). With poor training data, the overall “accuracy” could approach 100% but be completely wrong. Please rephrase.

Line 422: While the use of “Global South” and “Global North” is quite common, some see these terms as problematic as they are geographically inaccurate, deterministic, and paternalistic. If authors mean “lower wealth” they could just say that.

Technical corrections

None