Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., referee comment RC1
https://doi.org/10.5194/gmd-2022-85-RC1, 2022

Comment on gmd-2022-85

Anonymous Referee #1

Referee comment on "SHAFTS (v2022.3): a deep-learning-based Python package for simultaneous extraction of building height and footprint from sentinel imagery" by Ruidong Li et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2022-85-RC1, 2022

This study develops a deep-learning (DL) based Python package-SHAFTS to extract 3D building information (average building height and footprint) from publicly available satellite imagery. Compared to conventional machine learning-based models and single-task DL models, the proposed multi-task DL models can effectively improve the prediction accuracy. This study involves the fusion of multi-source input data and many machine learning and deep learning models, which undoubtedly requires huge and solid work from the authors. Although I am not the expert in computer science, the evaluation framework presented in Section 3 is scientifically sound from my perspective – very quantitative from patch-level to city level. And I will consider using the developed package in the future. I only have the following minor comments.

Minor comments:

- Line 368, the caption of Figure 4, need to denote the source of reference values when calculating RMSE, MAE, etc.

- Figure 5, is the density of scatter points normalized to $[10^{-4}, 1]$ instead of $[10^{-3}, 1]$?

- Line 385, this paragraph shall better describe the stratified error, which is helpful to the readers who are not familiar with it.

- Line 404-405, "Both SVR and DL models show relatively unfavourable RMSE and NMAD for the low-value domain but DL models behave slightly better." What does it mean? I think the RMSE and NMAD are both smaller in the low-value domain (Fig. 8).

- Line 463, what is contextual information?

- 12 and 13, why the feature patterns of STDL and MTDL have large differences?

- Line 509, I think STDL model gives better predictions than others from Fig. 15?