

## Comment on **essd-2022-259**

Anonymous Referee #1

---

Referee comment on "An open-source automatic survey of green roofs in London using segmentation of aerial imagery" by Charles H. Simpson et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2022-259-RC1>, 2022

---

### General comments

This is a detailed, well-written manuscript that distinctly describes the new vectorised green roofs dataset and the deep-learning method applied to automatically detect green roofs in visible aerial imagery. It is easy to gain access to the dataset and supporting code. The methods section provides useful information pertaining to hyperparameter tuning which is important for reproducibility. The results provide valuable insight into relative green roof coverage across London boroughs (e.g., Figures 4 and 5). Although this information is already contained within the datasets referenced and compared to within the manuscript (LRW2019 and London Plan AMR 16), it is evident this is the first time this data has been provided in an open access format.

My comments are primarily minor or typographical in nature, apart from one major concern. This relates to the lack of evidence that the trained U-Net can suitably generalise to other locations or other imagery of London. One of the benefits of training a neural network is the ability to apply the tool to automatically detect the feature(s) of interest in new images. Therefore, it is a concern if it is not possible to do this. The trained U-Net was not tested on imagery captured during different years or seasons in the year when lighting conditions may alter the appearance of the green roofs in the imagery. It is highlighted within the discussion section that the trained U-Net produced a lot of false positive results (over-predicted green roof coverage) in some Eastern boroughs of London. This is attributed to the potential use of a different collection instrument, highlighting that the trained U-Net may not be able to generalise to imagery captured using different sensors. It is suggested that the U-Net is applied to imagery captured during different years to test the model's ability to generalise. The imagery does not nearly need to cover the whole of London but cover one or two study locations to demonstrate the ability or otherwise for the trained U-Net to detect green roofs in a variety of settings.

Below are some other minor and typographical comments:

### **Specific comments**

Line 105- Please detail the number of images mosaiced to cover the whole London study area and the time period covering the first and last image. Also, is there information on the time of day when the images were captured?

Section 2.3- Was any pre-processing conducted e.g., pixel value normalisation?

Section 2.3- Please provide a little more information on how the trained U-Net was applied to unseen images. E.g., were images tiled or patched prior to input? If so, how were outputs mosaiced and were there any issues with predictions at the edge of the tiles?

Section 2.3- How was it defined whether a pixel was positive or negative? Was a threshold value applied to the U-Net outputs and was a consistent value applied across images/ image patches?

Line 169- was class imbalance considered within the loss functions? If so, how?

Figure A2- The orange and blue labels are the wrong way round in the legend and image description.

Figure A3- The orange and blue labels are the wrong way round in the legend and image description. There are also no green lines in the image.

### **Technical corrections:**

Line 86- change 'an' to 'a'

115- missing word between 'hand-labelled' and 'are'. Should it be polygons/ datapoints?

152- 'm' should be italics.

Throughout- be consistent in use of 'hyperparameter' or 'hyper-parameter'.

Figure 6- please add in small-scale schematic of London to show where the example images are from.