

Comment on **essd-2021-62**

Anonymous Referee #2

Referee comment on "The dataset of walled cities and urban extent in late imperial China in the 15th–19th centuries" by Qiaofeng Xue et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-62-RC2>, 2021

Records of urban extent are useful for not only investigating the urban development, especially for China, which has a long history of urban activate as well as among one of the fastest urban development hotspots worldwide. Many urban extent datasets have been produced in the past decades, benefited by the unprecedented capability of both Earth observation and machine learning; however, these datasets mainly focus on delineating the urban-related land covers, e.g., impervious surface, for the past 4 decades or so, due to the relatively short history of Earth observation. Extending the records deeper into the history of urban extent would provide a more complete picture of the urban development. Thus, I believe that datasets with an older/longer history of urban development would provide valuable information for complementing the urban records provided by the other urban land cover dataset. However, I think there are still flaws in both the manuscript and the presented dataset. These need to be addressed before the manuscript to be published.

1. The authors did not provide a clear definition of "urban" in the manuscript, which is commonly considered as "a primary nexus of human and environmental system interactions". The authors claimed to provide an urban extent dataset, but what it actually providing is the extent of the city wall. The definition gaps between the city wall and urban extent are considerable. Although the authors explain the relationship between these two, the connection is still too weak, and further clarifications would be needed. Can a city wall equal or able to represent the actual urban extent? The construction of the city wall would have taken decades to complete, and the extent of the wall could considerably lag behind the changes of urban extent. I can agree that the main urban activities were inside the city wall, but I don't believe it's completely within it, especially during peaceful and economically prosperous times. Both Ming and Qing dynasties had large territories and enjoyed long peaceful and prosperous periods, particularly for areas away from the border. Representing the urban extent with only the extent of city wall would significantly underestimate the actual urban extents for many cities in the two dynasties. The authors argued in the manuscript that the city wall has been used to study the extent of cities, and I think it could be reliable for some individual cities, but I could not agree that the city wall can be used as the indicator for all cities.

2. The cities wall were mainly built for large cities. Could it lack representation for smaller cities or towns?

3. The dataset has no validation. I understand that validation for such a historical record is difficult, but I am afraid that it cannot be accepted without comprehensive validation. Nevertheless, I think there could be still ways to assess its quality, for example, evaluating the changes of the extents for particular regions with reliable urban records?

4. Line 241-248, the authors explained the challenges of using the city wall for representing the urban extent. However, the manuscript did not carry out a convincing solution to address those.

5. Line 38, "did not start slowly". Please remove "slowly".

6. Line 150-151, why a city is considered a new city when it is chosen as an administrative center? What if the city was already there, and got chosen later to become an administrative center?

7. Line 166-168, The sentence does not make sense, please rephrase.

8. Line 226, why transform the data? Transform from what to what? Please provide more details.

9. Please clarify why the representative years were so precisely selected? How did the authors make sure the years, for example, 1648, and how all the cities had updated records for the exact year?

10. Line 202-203, Google Earth is a tool, not a platform.

11. Line 256, what's the implication of the correlation?

12. Line 268, please clarify how the proportion of cities were calculated? What's the number of total cities?

13. Line 296-297, please rephrase the sentence to fix grammar errors.

14. Line 298, add "a" before "slow rate".

15. Line 299, it would be odd to describe years using minimum and maximum.

16. Line 300-302, the periods were already explained in the previous sections.

17. Line 304 and 311, is the building of military cities actually related to urban development?

18. Line 324-325, the regions have long development but the city walls did not expand. Does it mean the city wall lagged behind while the urban experienced development?