

Atmos. Meas. Tech. Discuss., referee comment RC1  
<https://doi.org/10.5194/amt-2021-301-RC1>, 2021  
© Author(s) 2021. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Comment on amt-2021-301

Anonymous Referee #3

---

Referee comment on "A high-resolution monitoring approach of canopy urban heat island using a random forest model and multi-platform observations" by Shihan Chen et al., Atmos. Meas. Tech. Discuss., <https://doi.org/10.5194/amt-2021-301-RC1>, 2021

---

This study aims to integrate remote sensing data with ground-based temperature measurement to predict canopy heat island at a large spatial extent. Overall the idea is interesting, but the data size is very limited and the method seems to be a typical approach. Below are my major comments:

1) Line 55, "meteorological measurements and high-density observations", it is not clear what the authors are referring to here. Does it mean rural station against microclimate observations? Please rephrase here.

2) Line 104, "'Nanjing's UHI was observed to be 0.5 in 2005". It is important to clarify how the 0.5 is calculated, as it does not contain any spatial variation, and differs from the results in this study.

3) Line 110, my major concern about this study is the usage of only 3 snapshots of the satellite measurement. Only 3 days are selected over a 5-yr period from 2013 to 2017. Such data availability is surprising low. On top of this, this performance of RF model is only marginal, with  $R^2$  about 0.5 in the cross-validation. How would the authors justify the potential or accuracy of the model in predicting the urban heat island for practical usage, if the model is to be extended to more days under complex weather conditions?

4) Line 113, what does 0.5 intervals mean?

5) Line 124, Does the AHF data vary diurnally and seasonally? If it does not, then the AHF data may be less meaningful to be incorporated into the RF model. Instead of the LULC map, I will suggest authors to add the AHF map here.

6) Line 150, a table summarizing the predictors with their sources and resolution used in the RF model will be very useful.

7) Line 180, Different variables are used on different days, and the optimal Mtry and Ntree also change substantially. This essentially means the built model can only apply for a specific day. Given that the study only focus on Nanjing under clear sky conditions, I am concerned about the applicability of the model. If we utilize this method to study CUHI in the future, that means one will need to run it for every hour, and the accuracy is not guaranteed even under the clear sky condition. What I suggest may be a lot of work, but I think using more models and comparing their performance, and come up with a consistently well-perform model is really needed to enhance this paper to a higher quality/level.

8) Line 197, from my perspective, using a buffer size of 2-km to predict temperature at 30-m resolution is not scientifically sound. As expected, the estimated spatial variability of air temperature/CUHI is small. This contradicts the local climate zone framework that local urban landscape may dominate the air temperature under a similar weather condition. Did the authors check the spatial map of SUHI? I believe the spatial variability of LST will be much larger. I suggest authors add a map of LST and compare it with the estimated CUHI. That can help facilitate the discussions on your model.

9) Figure 6 and Figure S2, why is the number of stations different in three subplots?

10) Line 406. It is good that authors summarize the existing approach of predicting AT in the literature. Given that RF model is a very typical approach adopted by many studies, what is the novelty of this study? Maybe the authors can elaborate more here.