The manuscript “GHRSAT: the first global hourly dataset of all-sky remotely sensed estimates of surface air temperature” developed a hybrid method integrate random forest models and kriging techniques to estimate all-weather air temperature, and generate global all-weather air temperature products from 2011-2023. The proposed method improved the model accuracy compared to traditional RF algorithom. The content structure of the preprint is clear, the method is innovative and the topic is meaningful. My comments and questions for clarification can be found below.

Comments:

1. Line 120: The author uses zoning modeling and mentions that the zoning basis also refers to station density. I am confused that the model building effect depends largely on the representativeness of the sample. Is it reasonable to use low-density sites for modeling? In addition, will zoning modeling lead to boundary effects between regions? Why not build a global unified model?
2. Line 158: Air temperature is related to many factors. The variables input into the model in this article are only LST, NDVI, latitude and longitude, elevation and hour of a day. What is the basis for selecting these variables? Among them, only LST and hour of a day change over hours. Is the result mostly dependent on LST? Please show the feature importance of the models.
3. Line 255： In the validation part, the samples were randomly divided into ten parts, one of which was used to validate the model, which means that the training samples may include all sites, and there is no completely independent site for validation. What is the prediction accuracy of this method in non-site areas?
4. Line 289: The author developed 156 models for each region. The temporal variation of air temperature has certain regularities, and data from the same period in different years may provide effective information. Why does the author establish a separate model for each month in 2011-2023?
5. Section 4.2: The spatial validation in the preprint is based on the station scale, which cannot reflect the continuity of the generated product and the estimated effect of non-site areas. Please further prove it at the spatial scale.
6. The preprint lacks data cross-validation. For example, the air temperature estimated from geostationary satellites or reanalysis data all have hourly air temperatures. Please compare the with the published air temperature data or methods.