Qin et al. construct a 60-year (1961–2020) near-surface air temperature dataset over the glaciers of the Tibetan Plateau by fusing satellite and multi-source observations. The used ensemble learning model is described in detail and sufficient experiments are conducted to validate the reliability of constructed datasets. The manuscript is well organized, and all results are clearly presented.

Minor comments:
1) As the article is aimed at a data journal, I think it is better to include more key information on the dataset in the title of the manuscript, such as the spatial resolution (1km), temporal resolution (monthly) and used methods (data fusion or machine learning).
2) Section 3.2: It is necessary to present some reasons or your considerations for selecting the random forest model, by citing relevant literatures or adding concise discussions.
3) Eq. 8: Please give more description for DISO. How should readers interpret the value? Does a larger or smaller value mean better?
4) Please add legends for all scatterplots.
5) Whether the stations outside the TP in the Northern Hemisphere locate at glaciers? If yes, it is better to add a sub-figure in Figure 4 to show the validation results at such station.