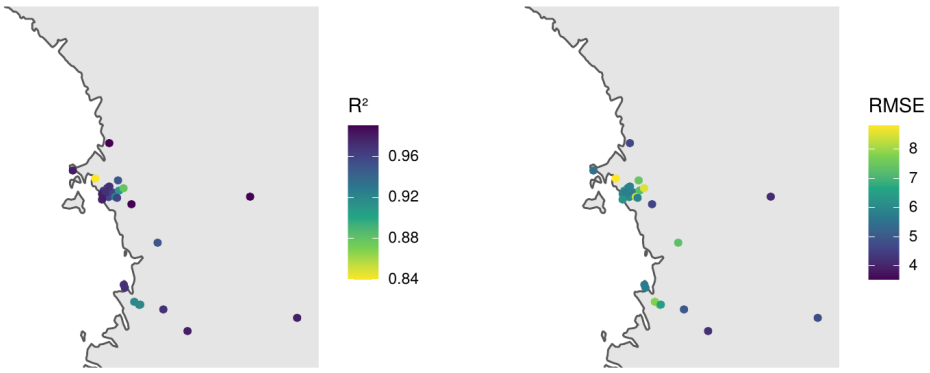
The presented dataset is able to reflect spatio-temporal patterns of air temperature in a high spatial and temporal resolution. We accept the concern that the accuracy could be further improved, and respect the decision to reject the dataset for this reason despite receiving a very supportive review from R1. However, we fundamentally disagree with the majority of the other points that were raised and it is important for us to express our opinion on a few points:

Apparently, a major concern of both R2 and the editor, is on the fact that biodiversity patterns are one (of several) motivations to develop this dataset ("*I found nothing responding to the "0.03%" affected land area claim by R2. Do authors have information to refute that number?*"). We're not refuting the percentage, and don't see the need to do so for our air temperature dataset. We explicitly state that the motivation of our dataset is intended for application in the context of several disciplines ("*Air temperature is an important baseline parameter for terrestrial Antarctica in the context of patterns and processes in climatology, hydrology or ecology.*"). Just because biodiversity is only limited to the ice-free areas does not mean that our air temperature product is limited to the same. The product itself, the figures given in the manuscript, as well as the text make clear that the dataset provides air temperature for entire Antarctica .

Both, the choice of the algorithm, as well as the validation strategy are described as casual by the editor and R2 respectively. We strongly disagree, as we outlined in our response to R2. Also, just as a minor remark, the point "*Meyer et al looked only at GBM*", which led the editor to the impression that the selection of algorithms is rather poorly justified, is not correctly cited in the review: already in the abstract of this paper, as well as in several other places, the cited paper emphasizes that "t*he performance of a simple linear regression model to predict Tair from LST was compared to the performance of three machine learning algorithms: Random Forest (RF), generalized boosted regression models (GBM) and Cubist.*"

For the Editor's comment on: "Also, according to their own Figure 1, worse RMSE where they have higher abundance of situ validation data? Does that not give them pause". This is not the case here. In the area with high density, some of the stations have rather high errors, but at the same time there are stations for which the error is very low (see additional figure below that focus on the area of the MDV where we have the highest density). This is very positive as we show that no general bias towards the areas with high density occurs.



Finally, the point that we very fundamentally disagree with, is that current trends and discussions on spatial predictive mapping developed in a context that might not be related to Antarctica is not of relevance here (*"They follow with a list of machine learning citations none of which mention Antarctica and none of which an Antarctic researcher will ever read."*). We are convinced that interdisciplinary collaboration is needed in the field of environmental monitoring and believe that this is the condition to make progress towards improved monitoring of environmental variables, both in Antarctica as well as elsewhere.

Finally, and despite our disagreement on some of the raised concerns, we would like to thank the reviewers and the editor for their work on our manuscript.