

Response to Referee #2 comments

The manuscript "A standardized permafrost ground temperature collection for Canada 2025" describes probably the most comprehensive compilation of standardized ground temperature time-series in permafrost areas across Canada. It includes data collected from repositories, but also extensive datasets provided directly by individual researchers and institutions. It includes both active layer and permafrost temperatures, as well as ground surface temperature. It integrates data from 29 published datasets and 13 unpublished contributions, encompassing 427 GT and 491 GST sites across British Columbia, Labrador, Northwest Territories, Nunavut, Québec, and Yukon.

Ground temperature data in permafrost regions are typically measured in complex and remote terrain, facing complex constraints, from borehole installation to data collection logistics, and even in data curation, and is many times still spread in different repositories, this dataset becomes especially relevant. It facilitates data collection for specific sites, but it also provides unvaluable data for the modeling community.

The data is presented in a standardized form into daily temperatures at specific depths, with the standardization process being described in detail. Due to the differences in the data sets, standardization needs to account for data limitations and this has been done by the authors in a clear and transparent way, which is well described in the manuscript. The data curation was supported by the software package tempcf which is available for community use and it works surprisingly well, allowing for changes associated to data cleaning to be recorded in a specific log. The dataset and all associated curation files are made available at Zenodo and easy to download. The same data is also accessible online via an ERDDAP server that allows for easy data visualization (including graphical) and download in multiple common formats. These are excellent practice and show the correct application of the FAIR principles.

Given the very high quality of the dataset, the clear organization of the data and curation of the data, as well as the completeness and clarity of the manuscript, I consider it should be accepted following minor reviews, which I indicate below.

Detailed comments:

- line 63 - consider changing "we also sent personal requests for contributions of unpublished ground temperature data to individuals" to "we also sent personal requests and received contributions of unpublished ground temperature data from individuals".

The sentence has been updated as suggested.

- Line 72: You mention that when not available, elevation data was collected from Google Maps Elevation API. Please clarify what is the data source that was used as DEM in that API, so that potential inaccuracies can be identified in the future (maybe the DEM quality will change with time). If this is not

possible to know, indicate this as a limitation.

Unfortunately, the data sources are not provided by the API. We have included the following text Section 4.5 (Sensor location) to acknowledge this limitation:

“Where elevation data were not available from primary sources, values were obtained using the Google Maps Elevation API. However, Google does not publicly disclose the underlying digital elevation model(s) used by this service, and the data sources may vary by geographic region and change over time as Google updates its mapping products. Consequently, the vertical accuracy of these elevation values cannot be independently verified or consistently reproduced in future studies. This represents a limitation of the present dataset, and users should be aware that any replication of this workflow at a later date may yield slightly different elevation values for the same locations.”

- lines 90-91: *“Measurements that clearly diverged from the surrounding ground temperature signal were removed.” Clarify what you mean with surrounding. I think it relates to time in the dataset and not to space. Please rephrase for clarity.*

We have rephrased it to: *“For a given time period, values diverging from the normal pattern established by other sensors within the same borehole were removed.”*

- Table 1: *Please include a column with the period of observations of each dataset. This will make it easier for readers to understand what periods are available.*

A new column with heading “Temporal extent” has been added to Table 1. It provides the year of first- and last- observations for each contribution.

- Figure 3: *Are the circles reflecting really the number of sites within 100x100km? They seem too close to each other for the area indicated for density. Please clarify.*

The markers represent the number of GT or GST sites within 100x100km grid cells displayed as the mean coordinate of all sites within the cell. The fixed nature of the grid cells results in some sites falling in different cells, even if less than 100 km apart. We plot the markers at their mean coordinates to best represent their placement in Canada, this is of particular importance for individual sites with smaller markers that would otherwise appear misaligned. We have shifted the grid cells to minimize these edge cases, but some still remain. The figure caption has also been revised to include the following sentence:

“The size of each point on the map reflects the number of sites within 100 km grid cells, displayed at the average location of the sites within that cell to improve visual clarity and location accuracy”.