1st submission

Editor Start Date: 3/26/2020

Editor Stop Date: 5/13/2020

Reviewer #1 (4/5/2020－5/4/2020)

Reviewer #2 (4/11/2020－5/2/2020)

Reviewer #1: Anonymous

The data paper entitled "Database for domestic ground temperature and freezing depth in Japan" contains of very valuable and informative long-term data on thermal state of whole Japanese stations. I highly impressed the digitizing work and thorough quality control to complying the dataset. I only show several minor comments as follows:

In the title of this dataset, I don't think there is a need to use the term "domestic".

The reader may not identify location of stations in Figure 1 in this form. I think location of stations on the map should represent using dots and the station number is put nearby the dot or using leader line. The number also should be adding white border to avoid overlapping with shoreline behind the text.

In Table 1, latitude and longitude of Nos, 10 and 11, Nos. 33 and 34 are same. Are these correct?

Reviewer #2: Anonymous

Comment to the manuscript #PDJ-D-20-00004 entitled "Database for domestic ground temperature and freezing depth in Japan" submitted to the Polar Data Journal

The paper describes the ground temperature data in Japan collected from different organizations and formatted with a common method. The reviewer believes that it provids valuable data for scientific research. The scope, method and contents of the data are clearly presented in general, but some of explanations, most of them are related to data sieving, should be added and revised. The reviewer recommends minor revision before publication. The followings are the points that should be reconsidered and/or revised and those that would be useful to improve the paper.

A. Points to be reconsidered and/or revised

A1. Abstract

The authors should mention that the meteorological data is included in the dataset, since that is an important and useful part to examine the ground temperatures and the freeze/melt depths.

A2. 2nd paragraph of "Data Records"

The authors describe the structure of the data files, but it is not clear what the terms in the file name, "product" and "full," indicate. That may be related to the unclearness about the convention of the meteorological data.

A3. Technical Validation

The reviewer understands from the 4th line that the flags A3-A5 are applied only to ground temperature. Did the authors intend that?

The lists of the sieving contexts (1.-4.) are helpful to the readers. The reviewer understands that the context 1. and 2. correspond to the flag A3 in Table 5 and the context 4. does the flag A4, but how about the context 3. and the flag A5?

Looking into detail, the temperature lower than -25℃ is taken too low in the context 1. However, the lowest record of temperature in Japan is lower than -35℃ (https://www.data.jma.go.jp/obd/stats/etrn/view/rankall.php). Taking that record into consideration, should the context criteria be as low as -40℃?

Similarly, the highest record of temperature in Japan is as high as 35-40℃. In addition, ground surface temperature becomes much higher than surface air temperature. Taking those into consideration, should the criteria of the context 2. be as high as 60℃?

Besides, the reviewer does not understand how the context 2. and 3. are distinguished in sieving, and why those two should be considered separately.

A4. Figures

The numbers of location are hard to see in the map.

A5. Tables

Please explain the meaning of "Rec#" in Table 1.

The periods of data or observation in Table 1 and Table 2 are different at some stations. The reviewer recommends distinguishing the observation period and the data period archived in this dataset, since they would be different at some stations.

The 2nd row of Table 2 contains multiple elements. The reviewer recommends separating each element in the separated rows, such as item, period, observation frequency, notes, etc.

The methods of "Best estimate" for the flag A1-A5 in Table 5 are not explained anywhere in the paper. Besides, the treatment of the data-filling failure for R5 is not explained as well. The reviewer recommends adding explanations for those in Section 3. (Methods), 5. (Technical Validation), or annotation of Table 5.

A5. Acknowledgements

Please check if "Dr. Yohitaka Muraji" would be "Dr. Yoshitaka Muraji.".

B. Points that would be useful to improve the paper

B1. 1st paragraph of "Background and Summary"

The logic in the beginning of the paragraph is not clear. Following is an example for revision.

"Subsurface thermal states (ground temperature and frozen or thawed ground) are important not only as climate indices, but also as information for activities in socioeconomic domains, such as agriculture and civil engineering. These states are thus important information not only in high-altitude or -latitude regions but also in Japanese domestic areas extending from the sub-tropics to the sub-arctic."

The authors wrote in the last sentence, "Past records of ground temperatures can provide basic information and understanding of how domestic subsurface thermal states have changed," but past records of ground temperatures can also provide information of above surface thermal states.

It is preferable that the authors would provide references for the statements of the 1st paragraph.

B2. 1st paragraph of "Methods"

The authors describe the ground temperature observation and data conducted by JMA in the 1st paragraph. It is preferable to summarize it in a Table (e.g., Table 2a) similar to Table 2 for those by HRO/ARD.

Authors Response:

Response to reviewer #1;

We thank the reviewer for providing instructive and sound feedback and suggestions. These have helped us improve the manuscript. Our responses to the reviewer's valuable comments and suggestions are systematically provided herewith.

1-1. In the title of this dataset, I don't think there is a need to use the term "domestic".

We agree that the use of both "domestic" and "in Japan" is repetitive. We have thus deleted the term "domestic" from the title.

1-2. The reader may not identify location of stations in Figure 1 in this form. I think location of stations on the map should represent using dots and the station number is put nearby the dot or using leader line. The number also should be adding white border to avoid overlapping with shoreline behind the text.

We have revised the maps in Figure 1 to improve visibility.

1-3. In Table 1, latitude and longitude of Nos, 10 and 11, Nos. 33 and 34 are same. Are these correct?

We have provided an elaborate explanation of the contents of Table 1 in response to the both the reviewers' comments. We have checked the relevant information including latitude and longitude, as well as distinguished the periods of observations (OY1, OYN) from the digitized data (DY1, DYN) in the table. We have also added an explanation of the entry items in the "Location" section.

Response to reviewer #2;

We are grateful to the reviewer for appreciating our efforts and making sound inquisitions and comments that helped us enhance the scientific and expressive contents of the manuscript. We have systematically addressed their comments and suggestions below:

A. Points to be reconsidered and/or revised

A1. Abstract

The authors should mention that the meteorological data is included in the dataset, since that is an important and useful part to examine the ground temperatures and the freeze/melt depths.

In response to the reviewer's observation, we have mentioned that the meteorological data are

included in the dataset.

A2. 2nd paragraph of "Data Records"

The authors describe the structure of the data files, but it is not clear what the terms in the file name, "product" and "full," indicate. That may be related to the unclearness about the convention of the meteorological data.

We have revised the file name conventions, and made the following revision to the text in the "Data Records" section:

One comma-separated-value-formatted file was prepared for each station and labeled "GTset_XXX_YYY.csv" where XXX and YYY denote the name of the organization (JMA or HRO in this dataset) and the station, respectively. The JMA files only contain ground temperature data, while the HRO files contain ground temperature data and the accompanying meteorological observation data when and where available.

Two additional files were included in the dataset. One file (GTmonthly_JMA.csv) contained monthly mean ground temperature data at the JMA stations compiled by the Ministry of Agriculture, Forestry and Fisheries and the JMA (1982)1. The other (airTdaily_JMA.csv) contained daily near-surface air temperature data at selected JMA stations, obtained from JMA's past meteorological data archive (http://www.data.jma.go.jp/obd/stats/etrn/index.php).

A3. Technical Validation

The reviewer understands from the 4th line that the flags A3-A5 are applied only to ground temperature. Did the authors intend that?

We coordinated our available resources on technical validation to focus on ground temperature because it is the primary target of the dataset. We have clarified the text in the "Technical Validation" section, and made mention of flag "A5" in Table 5.

The lists of the sieving contexts (1.-4.) are helpful to the readers. The reviewer understands that the context 1. and 2. correspond to the flag A3 in Table 5 and the context 4. does the flag A4, but how about the context 3. and the flag A5?

Looking into detail, the temperature lower than -25℃ is taken too low in the context 1. However, the lowest record of temperature in Japan is lower than -35℃ (https://www.data.jma.go.jp/obd/stats/etrn/view/rankall.php). Taking that record into consideration, should the context criteria be as low as -40℃?

Similarly, the highest record of temperature in Japan is as high as 35-40℃. In addition, ground surface temperature becomes much higher than surface air temperature. Taking those into consideration, should the criteria of the context 2. be as high as 60℃?

Besides, the reviewer does not understand how the context 2. and 3. are distinguished in sieving, and why those two should be considered separately.

The purpose of the screening criteria is to detect possible hand-written or digitizing errors. Thus, although criteria (1) and (2) are meant to detect possible outliers (corresponding to A3), they are set in relatively looser terms (with lower threshold values) than the range with the record minima or maxima. As the reviewer correctly pointed out, criterion (4) corresponds to flag A4. However, criterion (3) is meant to detect the failure to re-convert the special "negative-value" convention employed in the earlier JMA routine (rather related to R1). Screening criteria (1)-(4), therefore, are neither directly proportionate to the flags nor mutually exclusive.

However, we realize that our original explanation failed to convey the above-mentioned details and our intention regarding the criteria. We have thus enhanced the explanation in the "Technical Validation" section.

A4. Figures

The numbers of location are hard to see in the map.

We have revised the maps in Figure 1 to improve visibility.

A5. Tables

(1) Please explain the meaning of "Rec#" in Table 1.

The periods of data or observation in Table 1 and Table 2 are different at some stations. The reviewer recommends distinguishing the observation period and the data period archived in this dataset, since they would be different at some stations.

We have provided an elaborate explanation of the contents of Table 1 to improve clarity in response to the comments made by both reviewers. We have made a distinction between the periods of the observations (OY1, OYN) and the digitized/archived data (DY1, DYN) in the table, in addition to checking the respective information. We have also added explanations on the entry items in the "Location" section.

(2) The 2nd row of Table 2 contains multiple elements. The reviewer recommends separating each element in the separated rows, such as item, period, observation frequency, notes, etc.

We have revised the layouts and compositions of Table 2 to enhance clarity and consistency.

(3) The methods of "Best estimate" for the flag A1-A5 in Table 5 are not explained anywhere in the paper. Besides, the treatment of the data-filling failure for R5 is not explained as well. The reviewer recommends adding explanations for those in Section 3. (Methods), 5. (Technical Validation), or annotation of Table 5.

We have replaced the words "How it is transcribed in data" with "best estimate in context" to reflect what R5 was subjected to, and added an explanation of "best estimate" in the "Technical Validation" section.

A6. Acknowledgements

Please check if "Dr. Yohitaka Muraji" would be "Dr. Yoshitaka Muraji.".

We have made the appropriate correction.

B. Points that would be useful to improve the paper

B1. 1st paragraph of "Background and Summary"

The logic in the beginning of the paragraph is not clear. Following is an example for revision.

"Subsurface thermal states (ground temperature and frozen or thawed ground) are important not only as climate indices, but also as information for activities in socioeconomic domains, such as agriculture and civil engineering. These states are thus important information not only in high-altitude or -latitude regions but also in Japanese domestic areas extending from the sub-tropics to the sub-arctic."

The authors wrote in the last sentence, "Past records of ground temperatures can provide basic information and understanding of how domestic subsurface thermal states have changed," but past records of ground temperatures can also provide information of above surface thermal states.

It is preferable that the authors would provide references for the statements of the 1st paragraph.

We have revised the first paragraph in the "Background and Summary" section and added references.

B2. 1st paragraph of "Methods"

The authors describe the ground temperature observation and data conducted by JMA in the 1st paragraph. It is preferable to summarize it in a Table (e.g., Table 2a) similar to Table 2 for those by HRO/ARD.

We have added a table (Table 2a) containing a description of the JMA ground temperature observation

practices.

2nd submission

Editor Start Date: 5/13/2020

Editor Stop Date: 6/14/2019

Editor Comments to the Author:

Thank you very much for sending the revised manuscript. As I confirmed that the revised manuscript properly responded to the comments from the two reviewers, I think the current manuscript is acceptable as it is. I will put this opinion on the editorial board.