Comment on essd-2022-167
Anonymous Referee #1

Referee comment on "A newly integrated ground temperature dataset of permafrost along the China-Russia crude oil pipeline route in Northeast China" by Guoyu Li et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2022-167-RC1, 2022

General Comments

The paper submitted by Li et al. describes a permafrost dataset compiled from measurements made along an oil pipeline route in NE China. The dataset includes information on ground temperatures, soil moisture along with results of geophysical surveys and meteorological data. Instrumented sites are located at various distances from the pipeline and are both on the right-of-way (ROW) and in undisturbed terrain off the ROW which facilitates characterization of impacts of pipeline characterization, similar to that done along other pipelines (e.g. Burgess and Smith 2003; Burgess et al. 2010; Johnson and Hegdel 2008). Such datasets and associated analysis are valuable because they can be utilized for model calibration and validation and also to evaluate mitigation techniques to reduce impacts and improve design of future infrastructure projects (e.g. Burgess et al. 2010). For these reasons, the paper and associated database should of interest to engineers and those involved in environmental impact assessment. However the manuscript does require a number of revisions before it can be accepted for publication.

As mentioned above, this is not the first time monitoring programs have been established along pipeline routes in permafrost regions and the data utilized to evaluation design performance, effectiveness of mitigation techniques as well as the lessons learned being used for design of other projects. For example, there are several papers and reports on two pipeline corridors in NW North America (Norman Wells to Zama (NWZ) oil pipeline NWT Canada and Alyeska oil pipeline in Alaska). These include for example Burgess et al. (2010); Burgess and Smith (2003); Smith and Burgess (2010); Croft et al. (2021); Mosely et al. (2021); Johnson and Hegdel (2008), as well as other reports referenced in these. Data compilations are available for the NWZ pipeline (Smith et al. 2004, 2008) and may also be available for the Alyeska oil pipeline. It would be useful to refer to these other studies in both the introduction to provide the rationale for compilation of similar databases and also when interpreting the results of data analysis.
The paper could be better organized and some important information should be provided. In particular, Section 2 (Study site and instrumentation) could be better written and include some information from Section 3 (Data description). It would be better to clearly describe the instrumentation used to measure each variable (i.e. ground temperature, moisture content, air temperature etc.) and use a table to indicate which variables are measured at each site/borehole. This would reduce repetition and enhance clarity. Some of the information in Section 3 also describes instrumentation and should be moved to Section 2 so that all methods appear in the same section.

Information on distance of boreholes from the pipe centre line for those on the ROW needs to be clearly provided. Also, the ROW width at each site should be provided as well as the distance from the ROW edge for boreholes that are in the undisturbed terrain off the ROW. This information is important as it provides information on the amount of disturbance there might be at each site. This information could be provided in a table or as site plans in the supplementary information.

Section 3 doesn’t really provide a clear description of the data in the dataset but rather provides some information on instrumentation as well as presentation of some of the data in graphic form and interpretation of the results. It would be useful to have a brief and clear description of the data. For example, are hourly, daily, annual mean values provided; if annual values provided, what is the time period (calendar or hydrologic year) and how do you deal with missing values; describe any processing done to ensure data quality and any derived parameters that might be included in the dataset. The presentation of the data and interpretation of results could be in a separate section.

A number of specific comments are provided below. Some of these are editorial while others are related to clarity of the content. I hope the authors find the comments helpful in preparing an improved manuscript that should be acceptable for publication.

**Specific Comments**

L19-20 – suggested revision: “In 2011 we initiated a....”

L21-24 – Not clear. Do you mean you compiled an integrated dataset on the ground thermal state utilizing the various types of data that are mentioned in the first part of the sentence. You should revise to be clear.

L24–26 – Are you referring to the undisturbed sites off the ROW? Revise to be clear. A revision is also suggested: “Results demonstrate that permafrost has warmed between 2011-2020 in the vicinity of SLLP, manifested as an increase in ground temperature at depth likely in response to climate change.”
L28 – Insert “an: between “as” and “insulation”

L37-38 – I suggest you remove the older references (ones between 1999 and 2008) and add more recent ones such as Etzelmuller et al. (2020); Zhao et al. (2020) and Liu et al. (2021). You could also cite the most recent BAMS State of the Climate sections on permafrost (i.e. Smith et al. 2021 and Noetzli et al. 2021).

L39 – You could delete the older Zhang et al. ref and add newer ones such as Burke et al. (2020) and other recent papers (you could also cite most recent IPCC reports of 6th Assessment).

L41 – I think you can delete the older references here (before 2010) as you have several recent ones.

L70-71 – Permafrost presence/absence is inferred or interpreted from the ERT results rather than “detected”. A revision is suggested: “... electrical resistivity tomography (ERT) surveys were used to determine the distribution of frozen and unfrozen ground in the vicinity of the CCOPs......ERT surveys to characterize freeze-thaw dynamics in....”

L68-74 – I don’t think you need this much detail here as it should be covered in the methods section (section 2). You could just say that boreholes were instrumented to measure ground temperatures and ERT surveys were used to delineate frozen and unfrozen ground.

L82 – replace “are” with “were”

L88-90 – Revision suggested “Between 1972 and 2017, MAAT increased at a ...” You can also delete “climate wetting”.

L91 – You can delete the last part of the sentence from “retrieved” onward. The additional information can be provided in the Table (or its caption) if you need to provide references to previous studies.

L92-93 – Suggested revision – “Permafrost is warm with mean annual ground temperature......ranging from -1.8 to -0.4°C. The reference may not be necessary as it seems to be a textbook, rather than a reference that specifically reports on MAGT in the region – perhaps one of the other papers in the list would be more suitable.
The XA site, located in a permafrost wetland, is the most northern site and has the lowest air temperature, while China has the highest air temperature and permafrost occurs in isolated patches.

Can't you just say it is ice-rich permafrost?

replace “built” with “installed”

You could just say when the hole was drilled and instrumentation installed. It isn’t clear why it is important to say that it was finished 6 months before installation of AWS. If it was a long drilling time and likely caused significant disturbance to the thermal regime just give the number of days of drilling and the implication of this.

How far off the ROW were these sites – this information should be provided.

Some of the detail could be reduced by putting some of the information in a table (or figure) such as the distance from the pipe, or depth of borehole. The table could also summarize all the instrumentation or surveys at a site, i.e. add this information to Table 3 for example. In this section you could just mention the instrumentation used to acquire information on the various parameters such as GT, soil moisture etc. You should also indicate how wide the ROW is at each site as this will give the reader an indication of the amount of the disturbance.

Don’t you mean the boreholes were drilled vertically? (rearrange sentence)

Weren’t thermistor cables used in other BHs?

Do you mean precision (resolution) rather accuracy here? Wouldn’t accuracy be more related to the temperature sensor?

The resolution/precision of instrumentation, including the dataloggers used should also be provided.

You could just give the maximum measurement depth in the table as the various measurement depths would be included in the database. Alternatively, you could include this additional information as a supplementary information table. You could then
include the additional information in the table suggested in an earlier comment (distance
from pipe, ROW width at each site, and whether BH is on or off ROW).

L153 – Section 3 - Data description – You seem to be confusing methodology with dataset
description. Much of this information including type of instrumentation used should be
included in Section 2 which describes methods including instrumentation. In some cases
such as Section 3.2 there is a repeat of the instrumentation description that was given in
Section 2. Section 3 also presents results, analysis and discussion so some consideration
should be given to revising the organization of sections.

L190 – Fig. 3b doesn’t clearly show the relationship between GT and latitude and whether
it is linear or not. MAGT at DZAA for each undisturbed site could be plotted vs latitude to
better show this.

L192-193 – Are you sure this isn’t an issue with the sensors?

L197-199 – A bit confusing as Fig. 4 shows seasonal variation.

L199-200 – Are you referring to the thermal offset here?

L201-204 – A couple of other things to consider: ALT also responds more to shorter-term
variations in air temperature compared to the deeper ground temperatures for which the
higher frequency variations are filtered out. ALT will therefore show more interannual
variability as shown in Fig. 5. The ground at JB is ice-rich, and melting of ground ice as
thaw progresses deeper into the ground can lead to surface subsidence and consolidation
of the unfrozen material. Changes in ALT over time can be relatively small as ice-rich
material thaws compared to sites where excess ice is negligible. See for example Nyland
et al. (2021), O’Neill et al. (2019).

L217-218 – The key thing is that insulation doesn’t prevent heat transfer but reduces it.

L219-221 – As well as the effect of the pipe, any ROW disturbance such as changes to the
surface, clearing of vegetation will also have an effect on the ground thermal regime.
These changes could also effect the ground thermal regime off ROW (lateral heat
transfer). See for example some of the publications on the Norman Wells pipeline
mentioned earlier and also Burgess and Smith (2003), Smith and Riseborough (2010).

L225 – It isn’t clear why this is an artificial permafrost table – is this the site with
thermosyphons? If so, it should be mentioned first.

L228-230 – Are you referring to surface settlement (subsidence) here?

L245 – Figure 8 – Consider reversing the colour scale and using blue for higher resistivity (colder ground – permafrost) and red for lower resistivity (warmer ground – unfrozen). This is more intuitive and has been done in other papers presenting ERT results. If some of the transect is off ROW, indicate the ROW edge. Was there a topographic survey done? – no change in ground elevation is shown.

L246 – There acronym TPCT isn’t really needed – just say thermosyphons in the rest of the paragraph.

L246-262 – The papers mentioned earlier on the Alaska Alyeska oil pipeline might be relevant here as the effect of thermosyphons are described.

L269 – Volumetric moisture content doesn’t really have units. (dimensionless) You could give as %.

References


