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-RC3, 2022

This paper presents an in-situ observational dataset on permafrost thermal regimes along the China-Russia Crude Oil Pipeline (CRCOP) route in northeast China, consisting of meteorological observations, soil temperature and soil moisture content data, and electrical resistivity tomography (ERT) data. The analysis of this dataset shows that the operation of CRCOP has had profound effects on the thermal state of the ground. The results are useful for better understanding the responses of permafrost to climate change and engineering activities. This dataset is valuable for evaluating the integrity of pipelines and the effectiveness of measures to mitigate the permafrost thaw, as well as for validating numerical models. The manuscript is overall well organized and written, but there are still some shortcomings that need to be addressed.

Specific comments

1. Please enhance the description of data processing (e.g. how do you deal with missing data and how these missing data affect the results?) in the revised manuscript. This will be important for full understanding of this dataset.

2. Warm-oil pipeline dissipates heat into the surrounding permafrost, resulting in thermal and physical disturbances to the pipeline right-of-way. These disturbances can compromise pipeline integrity and pose the potential risk of oil leakage. Therefore, in-situ permafrost monitoring has been made along several important pipeline routes (e.g. the Norman Wells pipeline and Trans-Alaska oil pipeline), where reliable first-hand data has been collected. I suggest the authors mention those important studies as background to this study.
3, Section 2. I found the locations of the sites are not accurate enough, rounded to two decimal places. The locations of boreholes and ERT profiles are not given in the manuscript, nor in the associated dataset. I suggest the author can provide accurate locations (at least four decimal places in unit degrees) for these mentioned locations.

4, Ground temperature was automatically collected by the dataloggers of RTB37a36V3 and CR3000 or measured manually with Fluke 87/89. Please provide a description of the errors occurring in these measurements.

5, The ROW widths are equal at each monitoring site? Please give a clear description.

6, Figure 3, seems problematic in the caption. According to the caption, column (a) indicates the active layer, but ground temperatures in XT in (a2-4) were measured below zero degrees for several consecutive years, which actually implies permafrost at this depth.

7, In Lines 110-111, please give exact timing for the borehole drilling.

8, Line 232 to 240, ERT results show that a talik formed around CRCOP I is much larger than that around CRCOP II. What is the reason?

9, Line 255-257, it’s difficult to directly observe the 1.5 m cooling range of two-phase closed thermosyphons and the 4 m lateral extent of thermal disturbance in Fig.9c and d, please clarify this point.

10, In Figure 8 and Appendix D, please use blue for higher resistivity and red for lower resistivity in ER images, like the color scheme in Figure 9. This is more intuitive for readers.

11, In Figure 10, please add the unit for soil volumetric liquid water content.