Reply to referee 2

Dear Vladimir Romanovsky,

We appreciate the time and effort spent on the revision of the manuscript "Modelling the evolution of permafrost temperatures and active layer thickness in King George Island, Antarctica, since 1950". The comments and suggestions were very well received, and we have subsequently revised and improved the text and work presented. In the following, we provide point-by point replies to all issues raised. The reviewer comment appears in bold font, our replies in normal font and changes to the text which will be implemented in the revised version of the manuscript are in italics.

On behalf of all authors,

Joana Baptista

Reply to the comments:

Line 39-40: Not clear. More explanation and a reference or references is/are needed here to explain better.

Reply: The sentence was rewritten for clarity regarding the factors influencing the differences between simulated and measured temperatures. Additionally, two references were added: one related to the application of the TTOP model (Obu et al., 2020), and another addressing the influence of cloudiness on MODIS LST data, which was also used in the modelling approach (Østby et al., 2014).

"This difference is attributed to factors such as heat advection from meltwater and rain - common in Maritime Antarctica but not account for in the model - or to cloudiness, which contaminated satellite observations of Land Surface Temperature (LST) employed as model forcing (Obu et al., 2020; Østby et al., 2014)."

Line 46-47: Not clear statement. The difference between these two models is that TTOP is an equilibrium semi-analytic model, but the CryoGrid is a transient numerical one.

Reply: The repeated use of the expression "model approach" was creating confusion in the paragraph. To improve clarity, the text was simplified:

"Both studies emphasized the necessity of incorporating detailed information on ground conditions, lithology, and moisture availability to improve TTOP estimations."

We maintained the description of the CryoGrid model in a separate section, and on line 34 of the preprint, we now include the sentence:

"Several studies have aimed to enhance the understanding of permafrost temperature by using statistical-empirical models, such as the Temperature at the Top of Permafrost (TTOP), across different spatial scales to produce temperature estimations for one moment in time."

Line 88: "Ground sensitivity analysis" is a confusing terminology. Better to call it "model calibration".

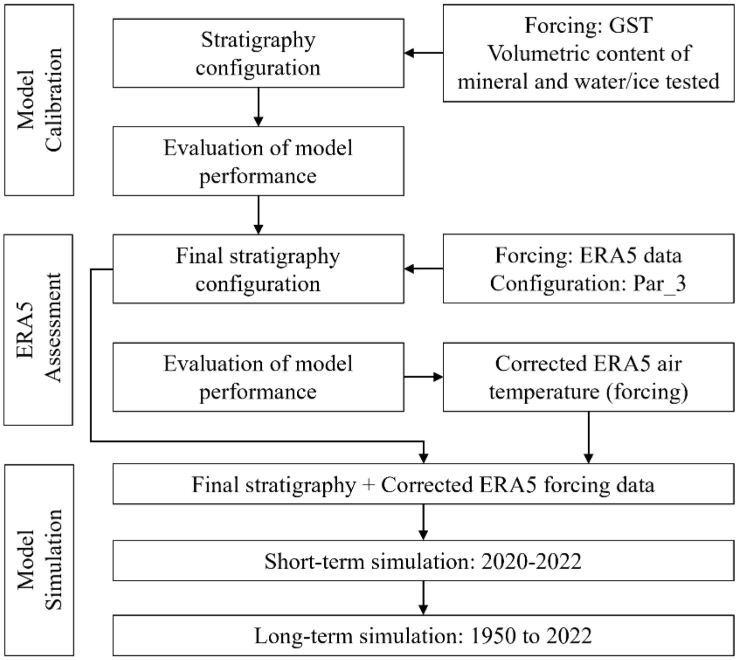
Reply: The sentence was rewritten for clarity. In the revised version, instead of using "ground sensitivity analysis" the term "sensitivity analysis to subsurface parameters" was used, since we are evaluating the simulation's performance accordingly with the changes on the subsurface parameters.

"In an initial step, the subsurface stratigraphy and associated thermal parameters were defined using a sensitivity analysis, for which simulations forced by the observatory's measured ground surface temperature (GST) were run with different combinations of parameters (referred to as "sensitivity analysis to subsurface parameters" hereafter)."

Figure 3: In the upper right block of text: Volumetric content of what and how is it tested?

Reply: We have changed the figure in the revised version of the manuscript. Additionally, Section 3.3.2 *Model Calibration* now includes an explanation of how the test was performed.

"For the sensitivity analysis to subsurface parameters, observed ground surface temperatures from January 2020 to December 2022 were used as upper boundary condition for the thermal model. The relative volumetric content of mineral and water/ice were tested in three combinations, with values ranging from 0.99 to 1 (mineral) and 0 to 0.02 (water and ice). For each simulation, the correlations between daily average ground temperatures at 0.02, 0.4, 1.2, 2, 3, 4, 6, 10 and 13 m depths with the observed temperatures were analysed, enabling to identify the parameterization that produced the best results."



Line 96: "…and integrated in the PERMANTAR…"?

Reply: The sentence was rewritten for clarity.

"The King Sejong Station (KSS) permafrost observatory, installed in March 2019, is part of the PERMANTAR network (University of Lisbon network of permafrost observatories in the Western Antarctic Peninsula)."

Line 105-106: Why monthly max, not the daily averaged temperature profiles?

Reply: We used maximum monthly temperatures to calculate active layer thickness because our objective was to compare annual variations between the observatory data and the simulation, rather than to capture high-resolution temporal dynamics.

"The Active Layer Thickness (ALT) at KSS was calculated by applying a logarithmic best fit to the maximum monthly temperatures at different depths allowing to identify the maximum depth of the 0 ºC isotherm."

Line 114-115: Not clear statement. Suggest more explanation.

Reply: We have added the following section after the first paragraph.

***"**It evolved from CryoGrid 1, which computes mean annual ground temperatures at the top of the permafrost (Gisnås et al., 2013). Subsequent versions introduced transient features for mapping temperature changes and a surface energy balance scheme for atmosphere-surface interaction (Westermann et al., 2023).*

The model version used here integrates functionalities from previous versions. Its modular structure combines classes representing distinctive surface conditions and ground columns, where their specific physics and state are accounted (Westermann et al., 2023). The objective is to stack and customise the classes that best represent the site for which the temperature is simulated."

Line 119: What is the depth of this column and what kind of lower boundary conditions were applied/prescribed?

Reply: The model has one-dimensional model column with a depth of 100 m. For the sensitivity analysis to subsurface parameters, we presented the results of the simulation for 13 m, corresponding to the depth of the borehole. The lower boundary condition was set to a geothermal heat flux. In the revised version, the sentence was rewritten as follows:

"For the sensitivity analysis to subsurface parameters, a one-dimensional model column, with 100 m depth was used, with temperature boundary condition (class "GROUND_freeW_ubT"). Measured GST was used as the upper boundary input, and a geothermal heat flux was applied as the lower boundary."

Line 133: "0.05 to 0.5 cm" Is it cm or m?

Reply: The unit is centimetres. The simulations were performed using a detailed grid resolution.

"The ground column was represented by the class "GROUND_freeW_seb_snow" in the CCM (Westermann et al., 2023), using a vertical grid resolution ranging from 0.05 to 0.5 cm and an albedo of 0.3 following Kim et al. (2006)."

Equation 2: in the right side of the equation (t,z) and z should not be subscripts

Reply: Yes. It was a typo that is now corrected in the revised version.

Line 144: 0.99 plus 0.02 makes 1.01, not 1 as it should be.

Reply: In this sentence, we refer to the range of values used in the combinations.

"The relative volumetric content of mineral and water/ice were tested in three combinations, with values ranging from 0.99 to 1 (mineral) and 0 to 0.02 (water and ice)."

Line 161-162: In this case is it logically to assume that the rest of the parameters (from a) through g)) will be also significantly deviated from the real ones? How can you correct this bias?

Reply: Among the remaining variables, in situ measurements were not available to directly assess potential biases. However, due to the issues encountered with the air temperature data, a literature review was conducted to evaluate the reliability of ERA5 in the Antarctic context. According to Tetzner et al. (2019), ERA5 shows a slight underestimation of wind speed in the coastal areas of the southern Antarctic Peninsula, likely due to the complex topography. Nonetheless, the same study highlights ERA5's general accuracy in representing the magnitude and variability of near-surface air temperature and wind regimes. Naakka et al. (2021) reported that ERA5 exhibits a warm bias in the interior of Antarctica during the winter months, accompanied by an underestimation of relative humidity. Despite these limitations, Graham et al. (2019) found ERA5 to provide the most accurate simulations of temperature, humidity, and wind profiles among five reanalysis datasets, including ERA-Interim, JRA-55, NCEP CFSR, and MERRA-2. Considering these findings, we acknowledge the inherent uncertainties in using reanalysis data. To address this, we corrected the deviations observed in air temperature and carefully accounted for the snow factor influencing snow depth. Since we aim to extend the simulations to the spatial scale, the lack of observational data, will require the use of ERA5 reanalysis data as forcing.

Line 178-179: "averaging 3.3 C for the season" – Is it air temperature or ground surface temperature? Cold season or mean annual?

Reply: It is air temperature. The sentence was rewritten to improve clarity.

"During the freezing season (FS) of 2020, intense ground surface cooling was observed due to lower air temperatures, that averaged -3.3 ºC, with daily lows reaching -20 ºC."

Line 181: “A similar pattern…” - What do you mean under "similar pattern"? Ground surface temperature during TS was higher in 2020/2021 than in 2021/2022, but it was colder for FS of 2020 compared to 2021 and 2022 freezing season.

Reply: The similar pattern is associated to the decrease of intensity in both seasons. In the revised version, the sentence has been removed to improve clarity.

Figure 5: Will be good to add a 0 C line to this Figure. It will help to understand the freezing/thawing process and the timing of it much easier.

Reply: Due to previous comments on the figure 5, we updated the display and now one line is visible at 0 ºC.

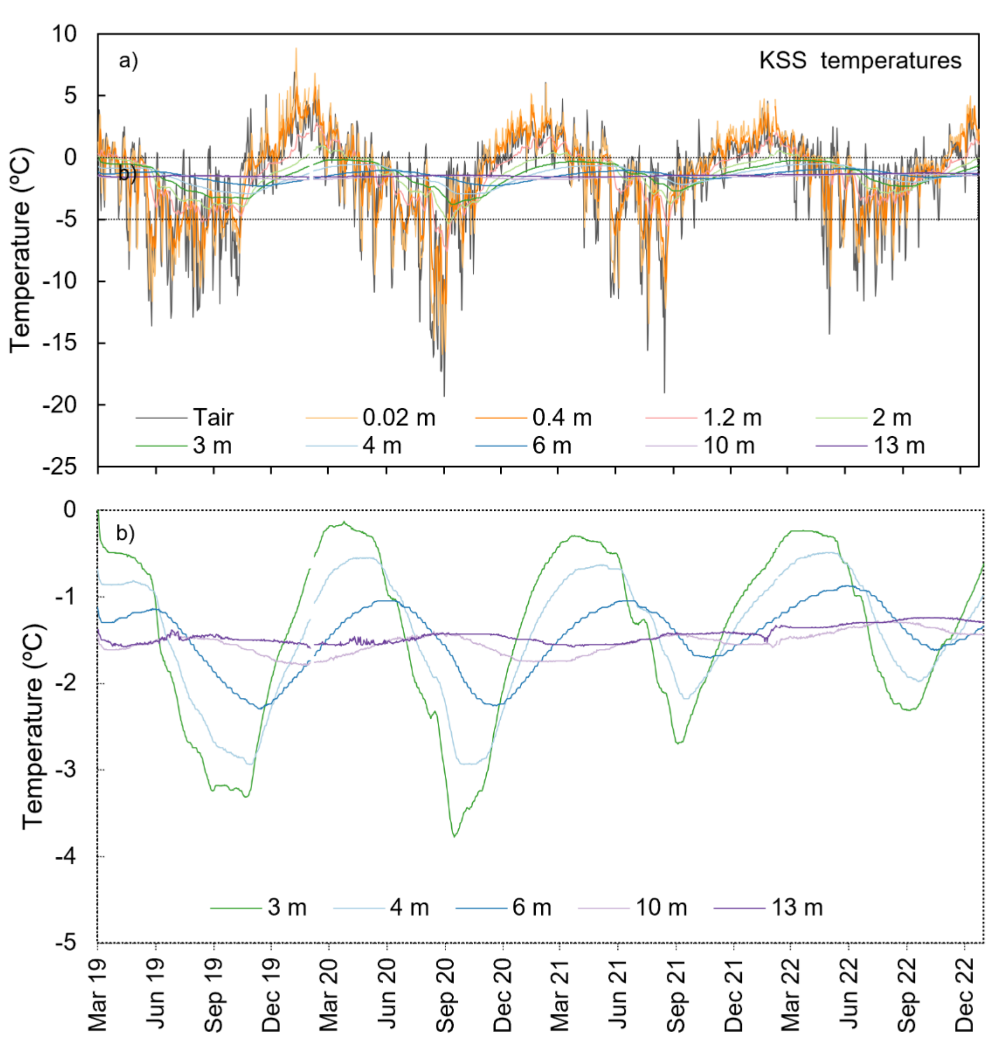
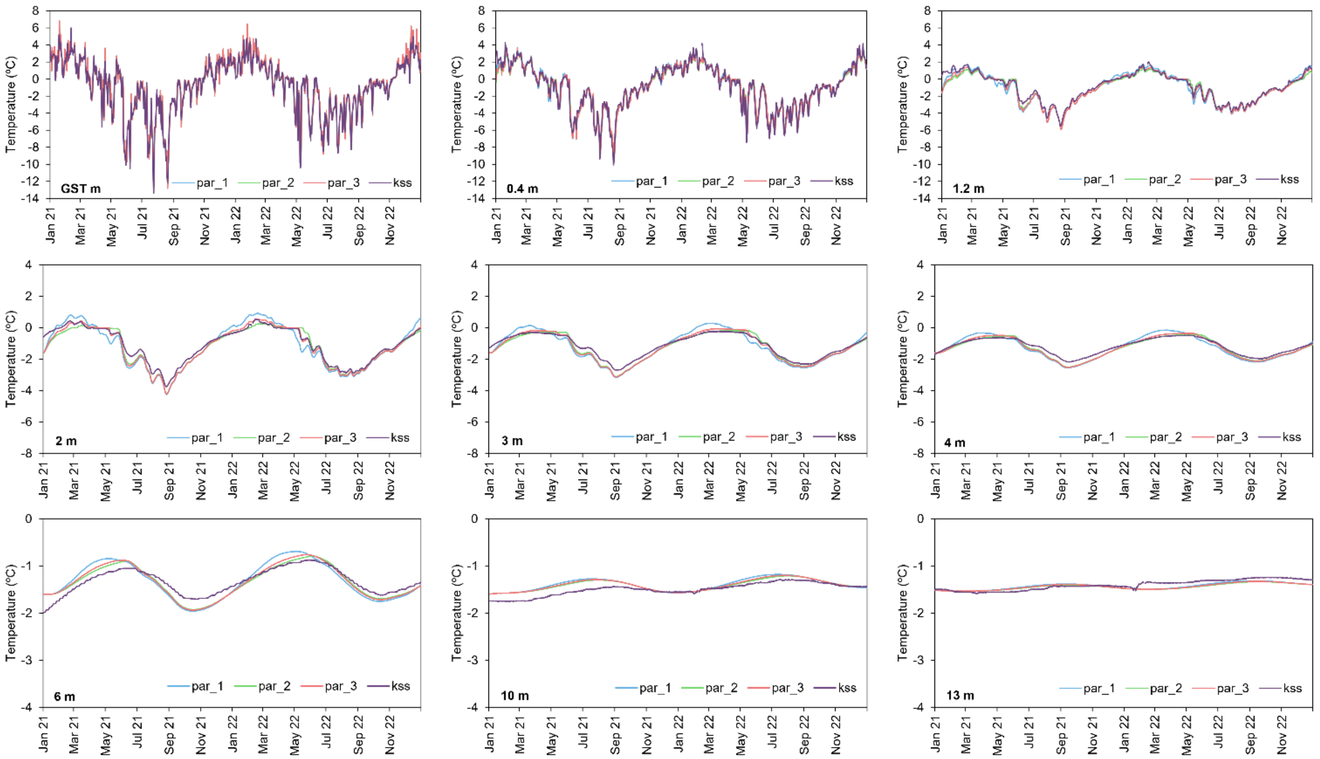
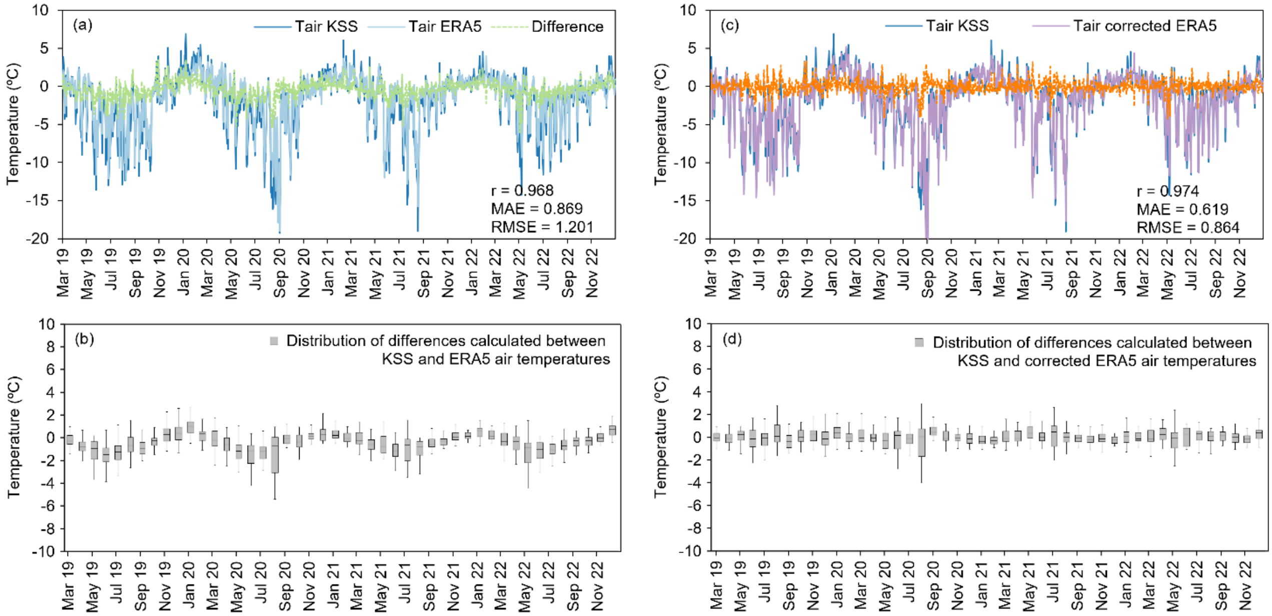
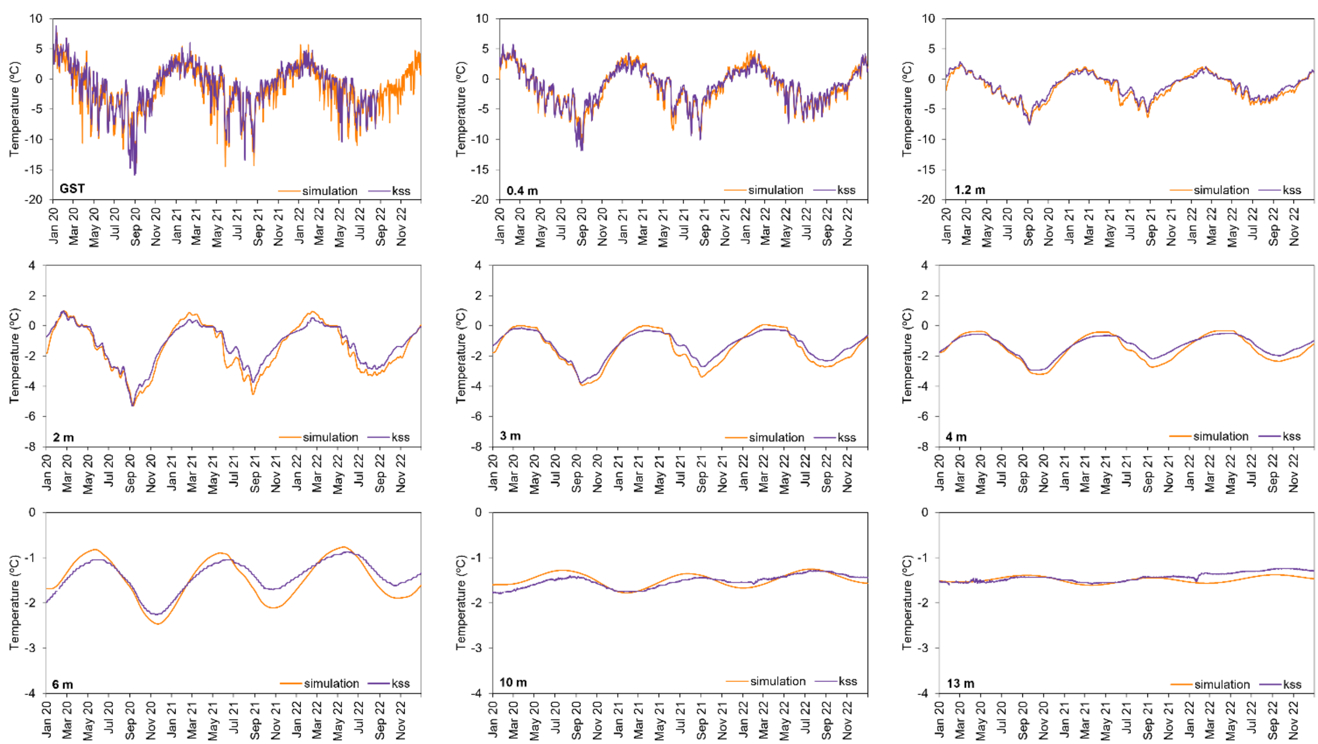


Figure 6-8: Very difficult to read numbers on the graphs. Suggest increasing the fonts.

Reply: The figures were updated.




(a)
Tair KSS
Tair ERA5
Difference
r = 0.968
MAE = 0.869
RMSE = 1.201
(c)
Tair KSS
Tair corrected ERA5
r = 0.974
MAE = 0.619
RMSE = 0.864
(b)
Distribution of differences calculated between KSS and ERA5 air temperatures
(d)
Distribution of differences calculated between KSS and corrected ERA5 air temperatures
Temperature (°C)


GST
0.4 m
1.2 m
2 m
3 m
4 m
6 m
10 m
13 m
simulation
kss
Temperature (°C)

Line 257-258: cannot see any evidence of the "zero curtain effect" in Figure 5. This Figure shows that the temperature in the active layer crossing the 0 C temperature threshold practically simultaneously at all depths. In this case we cannot talk about "zero curtain effect".

Reply: Figure 5 shows the overlap of ground temperatures at nine different depths, which limits the ability to observe the detailed dynamics of each line. However, in Figure 8, the thermal dynamics recorded in the borehole are presented in detail.

Line 278-279: Please, provide here or in Discussion section some explanation of this effect and why the model overestimates ALT.

Reply: On the section *5.2 Performance of CryoGrid simulations* we extended the discussion on the ALT overestimation.

"However, in Westermann et al. (2017), the ALT was calculated for sediments, while our site, featured bedrock. Due to its higher thermal conductivity and lower heat capacity, heat propagates more efficiently through bedrock than through sediments. As a result, during the thawing season, active layer depths are more sensitive to surface thermal forcing. Furthermore, this higher sensitivity means that even small misrepresentations of the surface boundary conditions - such as air temperature or incoming radiation—can amplify deviations in the simulated ALT."

Figure 9: How was this figure produced? It seems that it reflects the freezing process of the active layer not realistically. It shows that the freeze-up is happening form the bottom up. In contrast, Figure 5 shows that the temperature at the all observed depths in the active layer crosses the 0 C threshold practically at the same time.

Reply: Figure 9 shows the monthly maximum thaw depths calculated for both the observatory and the simulation. For the observatory, a logarithmic best fit was applied to the maximum monthly temperatures recorded by the borehole sensors, allowing the estimation of the maximum depth of the 0 °C isotherm. This approach was used to overcome the limitation imposed by the depth interval between sensors. For the simulations, due to the higher grid resolution, the ALT was determined by identifying the depth at which the maximum monthly temperatures exceed 0 °C.

Line 289: “Severity” is a strange choice of wording. "Severe warming" sounds a bit strange.

Reply: In the revised version, the word “severity” was replaced by “values”.

“During the thawing seasons, which lasted from November/December until February, the warming resulted in 241 and 198 TDDs in 2020 and 2021, respectively, indicating slightly lower values compared to the TDDs calculated for KSS, which ranged between 251 and 207 (Fig. 10 and Table 4).”

Line 290: We cannot see TDDs in Figure 10.

Reply: Figure 10 presents a contour plot showing ground temperatures at various depths over the analysis period. The TDD are shown in Table 4.

Table 4: Ground thermal parameters calculated for the KSS borehole between March 2019 and December 2022.

Parameter	KSS observatory			KSS simulation		
	2020	2021	2022	2020	2021	2022
MAGST	-1.7	-1.0	-0.9	-1.9	-1.5	-1.4
ALT	2.8	2.5	2.7	3.0	2.9	3.1
TTOP	-1.6	-1.1	-1.1	-1.7	-1.4	-1.3
MAGT 13 m	-1.6	-1.5	-1.3	-1.5	-1.5	-1.5
FDD	970	645	589	966	788	751
TDD	251	207	-	241	198	-

Line 311: My calculations show 0.51 C/decade for MAGST ((4.5-0.8)/7.2).

Reply: The MAGST for each year, from 1950 to 2022, was used to calculate the linear regression from which a slope of 0.0270 ºC was obtained. The average warming per decade was then calculated by multiplying the slope value by 10 years. Since MAGST values fluctuate due to colder and warmer periods, a linear regression equation calculated using all available values seems more suitable as it provides a better fit to our data.

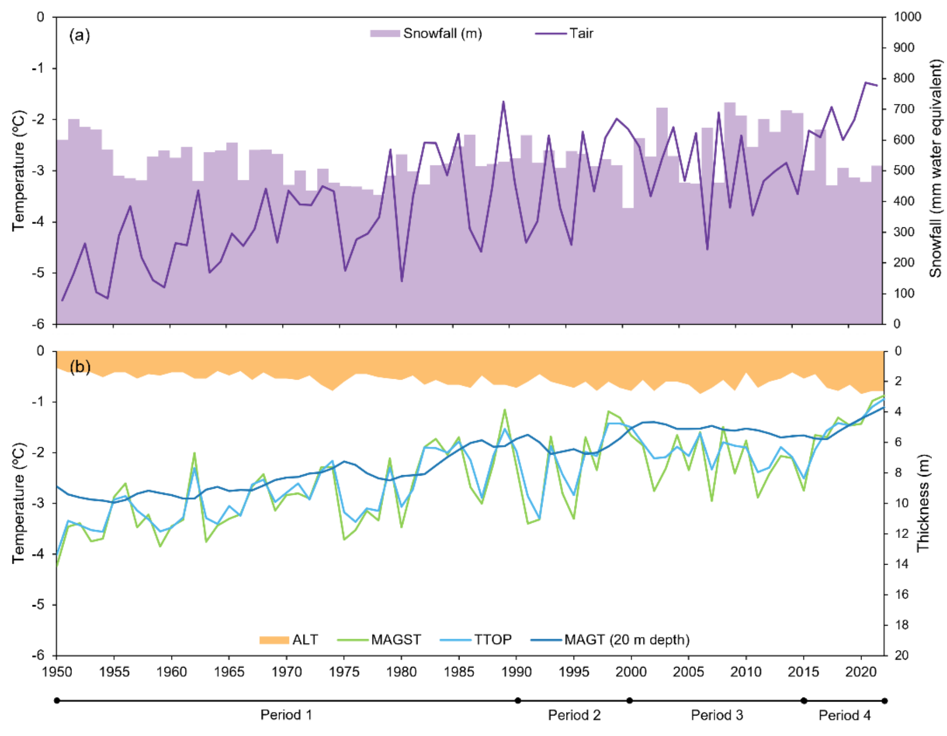
Line 312: Why is the geothermal gradient in the upper 20 m prescribed in the model for 1950 so large (0.09 C/m)? Also, my calculations of the warming rate at20 m depth show 0.21 C/decade.

Reply: The steady-state temperature profile was estimated through a model spin-up. Forcing data from ERA5 covering the period from 1940 to 1950 was used to run the spin-up 10 times in order to achieve an independent profile. We followed the approach described by Westermann et al. (2023), whom state: "While this state is usually reached within a few years in the uppermost meters, it takes much longer for deeper layers. For climate change simulations (…) a spin-up of several hundred years can be necessary, thus requiring significant additional computation time. However, reliable spin-up model forcing this long back in time is often not available, so that it is approximated by repeatedly looping a shorter period, often the first part of the regular model forcing."

Regarding the warming rate, the explanation was provided in answer to line 311. We are using the MAGT values over the entire period, not just those from 1950 and 2022.

Figure 12: More contrasting line colors will be very desirable.

Reply: We have changed the lines width to improve readability.



Line 327: Definitely wrong unit (probably should be cm or mm). Also, it needs to be said that it is snow water equivalent, not the snow fall

Reply: In the revised version, the values on the figure 12 were corrected and the text was updated.

"The analysis of the evolution of the permafrost and active layer temperatures, allowed for identifying four periods since 1950:

- *Period 1 (1950 – 1990): This period begins with the lowest MAGST and TTOP values of the series, around -4 ºC, and an ALT of 1.6 m. During the subsequent years, a progressive warming occurred with a rate of 0.04 ºC/year, following the increase of the air temperature (0.05 ºC/year). The snowfall shows the opposite trend, with a slight decrease of precipitation from 602 mm water equivalent in 1950 to 541 mm water equivalent in 1990."*

Line 328: ALT is controlled mostly by the summer conditions, not by MAGST or TTOP which are mean annual values.
Reply: In line 328 ALT variations are attributed to the "short periods of cooling and warming".

"MAGST and TTOP values show higher fluctuations due to short periods of cooling and warming, which control the ALT, that varied from 1.6 to 3.2 m. At 20 m, the temperature increased from -2.7 to -1.8 ºC."

Line 334: Please, check this number. It is not small and equal to the rate of warming during the Period 1

Reply: In line 334, no warming rate was mentioned. If the comment refers to the warming rate of MAGT, the following values were obtained: for the period 1 (1950–1990), a rate of 0.0273 was calculated, while for the second part of period 2 (1995–2000), a rate of 0.0885 was obtained.

Line 347: Please, find a better word here. "Severity" of the thawing season doesn't sound good for cold climate regions.

Reply: In the revised version, the word "severity" was replaced by "intensity".

"Concerning the freezing and thawing seasons, which contribute to the characterization of the four periods, a general trend of intensity decrease is observed for the freezing season, while the opposite trend is evident for the thawing season (Figure 13).

In the initial period (1950–1990), the interquartile range (IQR) for ground surface intensity was between 1090 to 1407 FDD. By contrast, in the most recent period (2016–2022), the intensity decreased, with an IQR ranging from 638 to 836 FDD, alongside with a narrowing of the range between minimum and maximum values. The freezing season duration showed no clear trend in the first three periods, with an IQR of 230–290 days, but decreased to 207–232 days in 2016–2022 (Fig. 13b).

In contrast, the thawing season demonstrates an increase in intensity. During the initial period (1950–1990), the IQR was between 137 to 222 TDD. In the final period (2016–2022), this range increased, with an IQR between 262 to 346 TDD (Fig. 13c). While thawing season duration showed no trend initially, with IQR spanning from 68 to 150 days, it extended to 138–154 days during the 2016–2022 period (Fig. 13d).”

Line 386: This fact was not explained in the Results section. May be it is a good place to speculate what was the reason in this overestimation of ALT in the model.

Reply: This is presented in Section 4.4*, Validation of Surface Energy Balance Model*, where the results of the comparison between the observed and simulated thaw depths are shown:

“The comparison between the observed and simulated thaw depths shows that the model forced with the ERA5 data, produced a slightly thicker active layer when compared to observations.

In 2021 and 2022, a maximum difference of 0.4 m was observed, with KSS ALT values of 2.5 m and 2.7 m, respectively, while the simulation predicted values of 2.9 m and 3.1 m (Fig. 9). In terms of the onset of thaw propagation, the simulation displayed a delay, with the active layer remaining unfrozen for up to one month longer than observed at KSS. This suggests a more pronounced deviation in the timing of ground thawing than in the intensity or depth of thaw propagation.”

Concerning the overestimation of ALT, in the revised version, we have addressed this paragraph in section *5.2 Performance of CryoGrid simulations*:

“Regarding the differences between measured and simulated thaw depth, the model predicted an active layer of 2.9 to 3.1 m, compared to observations of 2.5 to 2.9 m, resulting in a maximum difference of 0.4 m which is higher than the difference of 0.1 m obtained by Westermann et al. (2017) for the Lena River Delta. However, in Westermann et al. (2017), the ALT was calculated for sediments, while our site, featured bedrock. Due to its higher thermal conductivity and lower heat capacity, heat propagates more efficiently through bedrock than through sediments. As a result, during the thawing season, active layer depths are more sensitive to surface thermal forcing. Furthermore, this higher sensitivity means that even small misrepresentations of the surface boundary conditions - such as air temperature or incoming radiation—can amplify deviations in the simulated ALT.”

References:

Gisnås, K., Etzelmüller, B., Farbrot, H., Schuler, T. V., and Westermann, S.: CryoGRID 1.0: Permafrost Distribution in Norway estimated by a Spatial Numerical Model, Permafr Periglac Process, 24, 2–19, https://doi.org/10.1002/ppp.1765, 2013.

Graham, R. M., Hudson, S. R., and Maturilli, M.: Improved Performance of ERA5 in Arctic Gateway Relative to Four Global Atmospheric Reanalyses, Geophys Res Lett, 46, 6138–6147, https://doi.org/10.1029/2019GL082781, 2019.

Kim, J., Cho, H. K., Jung, Y. J., Lee, Y. G., and Lee, B. Y.: Surface Energy Balance at Sejong Station, King George Island, Antarctica, Atmosphere (Basel), 16, 111–124, 2006.

Naakka, T., Nygård, T., and Vihma, T.: Air Moisture Climatology and Related Physical Processes in the Antarctic on the Basis of ERA5 Reanalysis, J Clim, 34, 4463–4480, https://doi.org/10.1175/JCLI-D-20-0798.1, 2021.

Obu, J., Westermann, S., Vieira, G., Abramov, A., Balks, M., Bartsch, A., Hrbáček, F., Kääb, A., and Ramos, M.: Pan-Antarctic map of near-surface permafrost temperatures at 1 km2 scale, The Cryosphere Discussions, 1–38, https://doi.org/10.5194/tc-2019-148, 2020.

Østby, T. I., Schuler, T. V., and Westermann, S.: Severe cloud contamination of MODIS Land Surface Temperatures over an Arctic ice cap, Svalbard, Remote Sens Environ, 142, 95–102, https://doi.org/10.1016/j.rse.2013.11.005, 2014.

Tetzner, D., Thomas, E., and Allen, C.: A Validation of ERA5 Reanalysis Data in the Southern Antarctic Peninsula—Ellsworth Land Region, and Its Implications for Ice Core Studies, Geosciences (Basel), 9, 289, https://doi.org/10.3390/geosciences9070289, 2019.

Westermann, S., Peter, M., Langer, M., Schwamborn, G., Schirrmeister, L., Etzelmüller, B., and Boike, J.: Transient modeling of the ground thermal conditions using satellite data in the Lena River delta, Siberia, Cryosphere, 11, 1441–1463, https://doi.org/10.5194/tc-11-1441-2017, 2017.

Westermann, S., Ingeman-Nielsen, T., Scheer, J., Aalstad, K., Aga, J., Chaudhary, N., Etzelmüller, B., Filhol, S., Kääb, A., Renette, C., Schmidt, L. S., Schuler, T. V., Zweigel, R. B., Martin, L., Morard, S., Ben-Asher, M., Angelopoulos, M., Boike, J., Groenke, B., Miesner, F., Nitzbon, J., Overduin, P., Stuenzi, S. M., and Langer, M.: The CryoGrid community model (version 1.0) – a multi-physics toolbox for climate-driven simulations in the terrestrial cryosphere, Geosci Model Dev, 16, 2607–2647, https://doi.org/10.5194/gmd-16-2607-2023, 2023.