

## **Author Comments to REFEREE #1**

We thank the Referee for the constructive comments. Typing errors and grammatical mistakes outlined in the supplement were corrected in the text without further comments.

### **MINOR ISSUES**

*RC: The introduction should be re-structured. Some paragraphs are out of context.*

The introduction was restructured trying to incorporate the suggestions.

*RC: Validation: Please provide the Nash-Sutcliffe (NS) values for your validation when starting the simulation in the year 1860. I don't understand why for the first validation of the model during S2 and S3 you use NS, and for the validation of the whole model experiment starting in 1860 you only compare MAGT<sub>meas</sub> to MAGT<sub>mod</sub>. Give NS values as well, and provide the ranges of the NS values (see comment in attachment, page 354).*

Nash-Sutcliffe values and corresponding ranges for the long-term validation are now provided in the text in addition to RMSE.

*RC: Historic temperature data (Page 348 and 349): Reformulate the paragraphs to make your procedure better understandable. What data do you have for what time period? How is the data interpolated (spatially) or extrapolated (temporally)? Your methods are unclear to me.*

This issue has been addressed during previous reviews. We have now tried to further clarify this point by modifying our description the paragraph with the goal of clarification. However, the reconstruction of the regional and historic air temperatures was not done as part of this study, but by Hanssen-Bauer (2005) and Hanssen-Bauer and Nordli (1998). We therefore consider that the full discussion of this method is beyond the scope of this manuscript and we need to refer to the above mentioned references for detailed reading. This method is widely applied by the Norwegian Meteorological Institute.

To answer the question on what data we have for what time period, this is stated in the text (P350 L3-5):

“The historic air temperature series used as input data for the modelling therefore consists of monthly values until 2008 and measured daily values for 2008–2011.”

The following sentence was added to the end of the paragraph to summarize the available input data:

“The following air temperature data series are therefore available as input data for the historic and future permafrost modelling studies: a time series of monthly means from 1860 until 2008, measured daily means from 2008 until 2011 and monthly means from 2012 until 2100.”

*RC: Language: Check the use of present and past in your manuscript. Check the logical order of your statements and paragraphs.*

Done

## SUPPLEMENT

- 346 L5 *RC: I think that this could be misunderstood, permafrost is present at BH4 and all the higher elevation boreholes, I guess. This is described in the next sentences, but to avoid confusion I would include the other permafrost BHs as well, or delete the first part of the sentence, i.e. start with: "The lower limit of permafrost along the instrumented slope was ..."*  
Paragraph was clarified and corrected as proposed.
- P348 L9 *RC: Provide units*  
Units added
- P348 L22 *RC: six temperature regions, in Norway?*  
The six different regional temperature regions were calculated for entire Norway. Additional information was added to the sentence
- P350 L15 *RC: for what time period?*  
For the correlation to Fokstugu weather station the observed air temperatures from S1 (2008/2009) and S2 (2009/2010). This information was added to the sentence.
- P352 L6-12 *RC: Why is the model not initialised with the reconstructed historic air temperature series? Is the initialization performed with daily temperature values?*  
For the initialization of the model a Fourier fit on the observed **daily** air temperatures from S1 was used and in addition set to the mean air temperature of 1860-1869. This is explained in the text in detail including an equation for the fit. We chose to use the mean of this 10 year period, to avoid to accidentally choose an abnormal year.  
  
We used the parameterised seasonal cycle from S1 for initialisation as this approach produces a continuous function fitted to the observed values, and has worked satisfactorily in other studies. Furthermore, the acceptable fit of observed and modelled GT for present conditions seem to justify the approach.
- 352 L20 *RC: Is the stepwise optimization procedure applied to avoid compensating effects, or the ranges of plausible values? Please make your point more comprehensible.*  
The comment by the reviewer was not entirely clear due to omission of the verb in the second part of the sentence. the context implies that the comment should have been: "Is the stepwise optimization procedure applied to avoid compensating effects, or TO EXPLORE the ranges of plausible values?"  
  
Our procedure was motivated to avoid compensating effects but not to explore the ranges of plausible values. In fact, we have predefined plausible search ranges from literature and look for improving the model performance within these ranges. We have expanded our description in the manuscript to make this point more comprehensible.
- P353 L6 *earlier in the text, this was c\_eff*

Changed to  $c_{\text{eff}}$

P354 L14-19

*Long time series model validation*

*RC: What about other locations than PACE?*

Fig. 7 also shows Juv-BH4 and Juv-BH1. A few boreholes had to be chosen for the illustration of the validation result in this figure. Juv-BH4 (Fig.7, b) and Juv-BH1 (Fig.7, b) were chosen because they are of completely different substrate, namely bedrock and block field. PACE was chosen due to the fact that there is 10 years of measured data down to a depth of even 100 m that can be compared to the modelled results (Fig. 7a)

*RC: How does the Nash-Sutcliffe value behave?*

Values were added to the paragraph. The ME shows lower values at depth ( $>0.6$ ) and higher values close to the ground surface ( $>0.9$ ).

*RC: How well is the model able to describe seasonal variations?*

Fig.7 (a) shows the seasonal behaviour of modelled and observed GTs at different depths at the PACE site for 10 years and also shows that the model is able to reproduce seasonal temperature cycles.

P355 L25

*RC: What about the location Jetta?*

The modelled GT changes at Jetta are in the same range of those modelled at Juvvasshøe due to their proximity and similarity in climate. The paragraph was clarified.

P356 L10

*RC: Does that have scientific importance? If yes, explain what we learn. I would not so much argue with statistical significance, but rather with relevance, i.e. quantify the trend in ALT increase during these periods.*

The inclusion of the Mann-Kendall test was suggested in one of the first reviews. It was included for that particular reason.

P359 L20

*RC: What are the values for Nash-Sutcliffe model efficiency coefficient? Mean values may compensate for seasonal effects.*

The Nash-Sutcliffe values were calculated between the “warmest” and “coldest” model run. Between these extreme value model runs the depth-averaged ME varies by less than 0.15. This was added to the paragraph.

P360 L6

*RC suggestion: “duration of the pronounced zero-curtain?”*

Changed in the text

P361 L19

*what would a significant deviation be? Can you quantify this?*

We agree that the word *significant* might be misused in that sentence. We have not performed a test on significance, so removed the word *significant* from the sentence.

P373 Table 1

*Why are some boreholes missing (JUV BH5 for example)?*

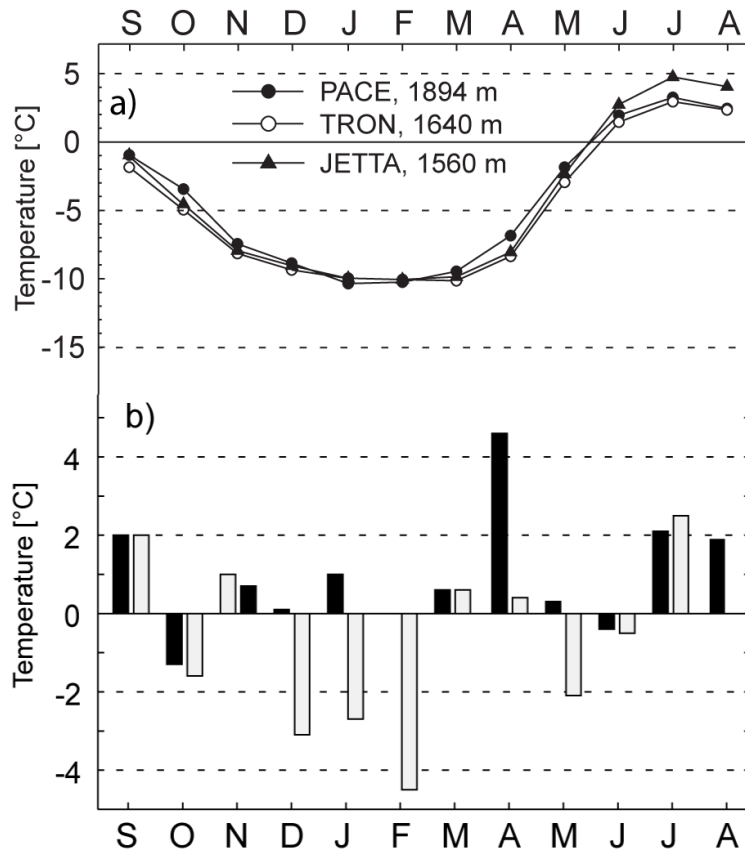
In the table only the boreholes that were modelled in this study were included to keep consistency and overview. This is now explained in the first paragraph of Chapter 2:

“Data from all boreholes were used in this chapter to give an introduction to climate setting and altitudinal variations at the study sites. However, only a selection of nine boreholes (Table 1) was used for the modelling study.”

See also reply to Referee 2.

P377 Fig. 2

Changed to



P384 Fig. 8

Scales adjusted and figure changed to:

