Reply to REFEREE #1, REFEREE #2 and REFEREE #3

We thank the Referees for the constructive comments to help improve the article. Obvious typing errors, grammatical mistakes and minor comments were corrected in the text without further comments. Corrections proposed in the supplement by Referee #1 were done without further comments.

Referee comments are in **bold**.

All three referees had the same or very similar comments on issues concerning the construction of air temperature data, the n-factor approach and calibration and validation process. Therefore, these comments are answered together for all referees in Part I. Individual replies to each Referee follows in the Part II.

PART I: Common Issues

1. Generation of Standardized Air Temperature Series and Historic Mean Monthly Air Temperatures

(SUBSECTION 3.4, page 822 line 9 - 21)

(Referring to Ref 1 / 1b + e; Ref.2; Ref. 3 / 3.)

What are standardized air temperature series?
How have they been generated?
How is the MMAT series created (Equation 2)?

The chapter has been reviewed and rewritten in order to explain more clearly how the historical air temperature series are created. However, describing the method in detail would exceed the scope of this paper. Therefore, detailed information about the method and data quality of the historic air temperature series can be found in detail in Hanssen-Bauer and Nordli (1998) as well as Hanssen-Bauer et al. (2005):

"Analyzing available longterm temperature records (starting in the 1860ies), Hanssen-Bauer (2005) and Hanssen-Bauer and Nordli (1998) identifiedsix temperature regions, each of which characterized by similar longterm variability of air temperature.
For each regionmonthly standardised series ST_m are derived by averaging the standardized temperature series $ST_{m,i}$ of each individual station i in the region m:

$$ST_m = (1/n) \times \sum_{i=1}^{n} ST_{m,i} \qquad (4)$$

The individual standardized series are presented as anomalies in terms of standard deviations $\sigma_{Tm,i}$ relative to the 1961-1990 average $\mu_{Tm,i}$ (Hanssen-Bauer, 2005):

$$ST_{m,i} = (T_{m,i} - \mu_{Tm,i})/\sigma_{Tm,i} \qquad (5)$$

where $T_{m,i}$ is the observed temperature series at station i in region m.
For the entire mainland Norway MDATs are available as 1km-resolution maps ($MDAT_{grid}$) for the period 01.09.1957 until today (provided by met.no, available at http://senorge.no, from hereon refered to as seNorge dataset). These grids are interpolated(kriging) from recorded temperatures at synoptic weather stations (Mohr, 2009). Daily air temperatures from 1957 to 2008 were generated for the boreholes PACE, Jet-LB1 and Tro-BH1 using linear regressions between measured temperatures and those extracted from seNorge for the corresponding location. This

procedure worked well for PACE with $r^2 = 0.8$ and a RMSE of 3.1 °C. For Jet-LB1 and Tro-BH1, however, the relation between observed air temperature and the corresponding seNorge value is non-linear, displaying a sharp bend at low temperatures. This characteristic is associated with the frequent occurrence of temperature inversions during winter (Farbort et al., 2011), a known drawback of the seNorge dataset which does not account for inversions during the interpolation procedure. To cope with this problem two separate linear regressions were performed for each site, one above and one below a threshold temperature (-10°C and -5°C for Tro-BH1 and Jet-BH1, respectively).

For the normal period 1961-1990 mean monthly values ($MAT_{i,1961\text{-}1990}$) and monthly standard deviations ($\sigma_{1961\text{-}1990}$) were calculated from these daily air temperatures. ST_m was used to construct a time series of monthly air temperatures at the station i (MAT_i) from the early 1860ies until today at the station i by (Hanssen-Bauer, 2005):

$$MAT_i = MAT_{1961-1990} + ST_m \times \sigma_{1961-1990} \qquad (6)$$

The observed temperature lapse rates during 2008-2010 of 0.5, 0.6 and 0.8 °C $100m^{-1}$ at Juvvass, Jetta and Tron, respectively (Farbrot et al., 2011), were used to transfer the so-constructed MAT_i time series to the other borehole locations (Table 4). The historic air temperature series used as input data for the modelling therefore consists of monthly values until 2008 and measured daily values for 2008–2010."

2. **For what time period is the MDAT_grid data available?**
   Why is the year 1957 chosen like kind of threshold?
   Why is it important, and why do you have to apply different methods before and after 1957?

   (Referring to Ref. 1 / 1a; Ref. 3 / 4.)

Due to the larger number of available weather stations for the interpolation after 1957, the gridded air temperature data across whole Norway is available starting from that point of time. Therefore the linear correlation between measured data at the borehole sites and the air temperature from the according grid cell is possible since that point of time.

3. **Uncertainty of and due to the n-factors determining the upper boundary conditions of the heat flux model. The authors should perform at least a simple uncertainty analysis using both n-factors, and some values in between, and discuss the differences in the results.**
   How can you demonstrate that air temperature and snow conditions observed in S1 and S2 represent the variability of the last decades?

   (Referring to Ref. 1 / 2b; Ref. 2 / Subsection 3.2., page 820 line 18-21)

We agree that the time period of only 2 years is short to derive n-factors that represent the whole modelled period. To give a better estimate on the representativeness of the chosen n-factors they were put in context to the 10-year time series and n-factors observed at PACE and Juv-BH5 (Isaksen et al., 2011, in press). The following paragraph was added to the discussion chapter 5.1 "Model uncertainties due to snow cover and water content variability":

"However, a 10-year record (1999 – 2009) of GST and T_{AIR} is available for the PACE and Juv-BH5 boreholes (Isaksen et al. 2011), which made it possible to give an estimate on the variation of n-factors and put the period 2008 – 2010 into context. At PACE a mean n_F-factor of 0.91 (0.89 – 0.98) and n_T factor of 1.12 (1.02 – 1.26) with standard deviations of 0.03 and 0.07, respectively,

was derived from the records. Juv-BH5 is characterised by larger variation of snow cover, resulting in mean n_F- and n_T-factors of 0.66 (0.29 – 0.93) and 1.15 (1.10 – 1.25) with standard deviations of 0.20 and 0.21, respectively. At the PACE borehole, logger locations of the 10-year period and the period 2008-2010 coincide exactly, so that the values can be compared directly. The mean n_F- and n_T-factors from 2008 – 2010 (Table 1), are well within the variation of the period 1999-2009."

A simple uncertainty analysis to estimate the error introduced by the n-factors is presented in this paper in Section 5.1. Running the model with n-factors from the previous or following season made it possible to give an estimation of the effect of n-factors. Varying the n-factor in the model simulates e.g. an increase or decrease of snow depth or a change in surface cover. As an example, a lower n_F-factor implies a stronger surface offset and thicker snow cover. A lot of different sensitivity studies using numerical modeling have already been done to investigate this effect in detail (Luetschg et al., 2008;Stieglitz et al., 2003;Zhang, 2005;Zhang et al., 2001). Further, a sensitivity study including varying n-factors using the same heat flow model was done by Etzelmüller et al. (2011, this issue). Therefore, further detailed investigation on that issue has not been included in this study.

4. Calibration and Validation

Calibration and validation periods consist of 1 year each (except at PACE). It is unclear how the model is calibrated and validated. Give more detail of the calibration procedure/methods used, and the time period for calibration/validation.

The calibration time period must be independent from the validation period. It is unclear whether the two periods were separated completely.

Time period of the data availability is very short for such large extrapolations in time, please discuss that further.

Can you provide more details on procedure of model calibration?

(Referring to Ref. 1 / 3. ; Ref. 3 / 2, page 821, Subsection 3.3))

The monitoring network on Juvvass, Jetta and Tron has been installed in summer 2008. Therefore, unfortunately, only 2 years of data were available for this study. In order to have independent calibration and validation periods we were forced to split this time period, so that one complete year was available for each. The periods used for calibration and validation coincides with the previously defined seasons. Therefore, the time period September 2008 – August 2009 (S1) was used for calibration and September 2009 – August 2010 (S2) for model validation. In that case the calibration and validation time periods were treated separately.

To make it more clear in the text Subsection 3.3. was devided into a new subsection *3.2. Model calibration* and *3.3. Model validation*. The following chapters were added to 3.2. and 3.3. to make calibration and validataion process more clear:

"The calibration was conducted by running the model for S1 with measured ground surface temperature as upper boundary condition and by adjusting the thermal properties of the subsurface (c, k, ρ and volumetric water content) in a stepwise, manual procedure until satisfying agreement between observed and modeled GTs was achieved. Values for thermal conductivity and density were measured at Juv-BH4 by the Norwegian Geological Survey (NGU) and served as initial guesses during the calibration. A time series of measured soil moisture in the vicinity of some sites (Juv-BH1, Tro-BH1) served as an estimate for the water content (O. Humlum, pers. Comm. 2011). Initial values for heat conductivity were taken from literature (Williams and Smith, 1989). Where

no direct measurements are available, parameter values were adjusted manually in order to achieve the greatest agreement between observed and modelled GTs. The main parameters significantly controlling the ground temperatures are the thermal conductivity and water content. Therefore, mainly these parameters were adjusted, however only small alterations to the initial values and literature values had to be done. Resulting values of the ground parameters used in the modell for the different materials are shown in Table 2.
Both in the calibration and validation process the Nash-Sutcliffe model efficiency coefficient (ME) was chosen to assess the agreement between observed and modelled GTs at each depth individually (Nash and Sutcliffe, 1970). Depth-averaged values of ME for each borehole are shown in Table 3."

"The data from S2 served as validation (V) period to assess the model accuracy. At the PACE borehole the time series was divided into a calibration period from 1999 to 2007 and a validation period from 2008 to 2010. A good qualitative agreement between measured and modelled GTs is presented in Fig. 6. The calibrated borehole models reproduce the measured GTs of the validation period with high accuracies showing ME-values ranging from 0.72 to 0.97 (exept for Tron BH1 with 0.54) (Table 3)."

.

PART II: Individual Comments

Reply to REFEREE 1

Major Revisions:

1. **Page 822, Section 3.4 (Historic and future temperature data), lines 9 - 23: This paragraph is very unclear. The procedure of generating historic air temperatures determines the main input factor to the heat conduction model to estimate the (past) permafrost conditions at all study sites. I am unable to understand how the historic air temperatures are generated, and can thus not judge whether the procedure is correct and of scientific use. Please be more precise:**

   see Part I, Nr. 1 + 2

2. **Model uncertainty is treated to a certain degree in the discussion. However, some of the largest uncertainties (only two years of calibration data and its influence on n-factors, estimation of historic air temperature) are not treated. Include a section concerning the assessment of model uncertainties in the methods chapter.**

   a. Discuss the uncertainty due to the short periods of measurements. How does this influence the results (model calibration, influence on n-factors and thus on the upper boundary condition)?

   The chapter on uncertainties related to n-factors derived from a short time period was extended (Chapter 5.1). Further analysis on the n-factors in relation to two 10-year time series has been done to put the n-factors of the period 2008-2010 in context (see Part I, Nr 3 and Chapter 5.1).

   b. see Part I, Nr. 3

   c. Treat the uncertainties arising from generating historic air temperature series. Uncertainties come from interpolation in space and time, linear regression, etc. This should at least be discussed qualitatively in the discussion, if quantification is not possible.

   The following chapter has been added to the discussion as 5.2.”Uncertainties due to the generation of historic air temperature series”:

   “The method by Hanssen-Bauer and Nordli (1998) has proven useful in reconstructing reliable air temperature time series (Farbrot and Hanssen-Bauer, 2009), however it introduces further uncertainties due to the spatial and temporal interpolation of air temperatures. Main uncertainties with the interpolation in mountain topography arise from uncertain lapse rates due to strong inversions (Tveito and Førland, 1999), which are occurring regularly at our study sites. Generally, the most extreme temperatures cause errors in the interpolation: cold temperatures are estimated as being too warm and warm temperatures as being too cold (Tveito and Førland, 1999). However, generally a good fit has been achieved when comparing measured and interpolated air temperature (Tveito and Førland, 1999).”

   d. Discuss the uncertainties due to the simple modeling approach, i.e. neglecting processes (advection for example).

   We are aware of these deficits. The following chapter was added to the text as Subsection 5.3 “Model uncertainties due to the simplified model approach”:

"In this study, we applied a simple 1D heat flux model, that neglects certain processes like advective heat transport. For our study sites insufficient measurements on water content and no information on lateral heat transfer are available. A direct estimation of the role of these processes and a validation of the model is therefore not possible. Further, 3D-effects altering the normal temperature distribution can not be represented accurately. Farbrot et al. (in press) indicates some 3D effects because of variable snow cover at the BH5 site. In our 1D modelling this effect is probably compensated by the calibration parameters, so that the performance of the model is relatively good. Hence, the results from BH5 in deeper soil layers has to be treated with caution.
Several studies have shown that heat conduction is the main process controlling the thermal regime of permafrost and advection plays a minor role (Weismüller et al., 2011)
Due to the following reasons we consider heat conduction as the main process of heat transfer at our study site:

- *Coarse grained material is present at all boreholes and thus, the water content is relatively low at all sites. This strongly reduces errors and modelling complications introduced by the advective water movement.*
- *The boreholes are drilled in flat topography. Processes of lateral heat transfer along a slope are negligible. Further, anomalies due to special air circulations within the block fields along a slope seem less important.*

Including all relevant processes in modeling is of course important for process understanding, and may have important effects on a short time scale. This study focuses more on a long time scale, and accordingly we assume conduction and latent heat effects as main factors here. However, we have not investigated whether processes other than those considered by the model would introduce a bias to our modeled GTs. Nevertheless, even with the stated simplifications, modeled GTs agree well with observations and the present borehole temperature distributions are reproduced when simulating the evolution since 1870. This suggests, that the simple model is capable of capturing the dominating processes."

e. What influence on the (historic) model results has the fact that during the validation period (2008 – 2010) the model is run with daily temperature values, and before 2008, with monthly data? The authors could quantify this by running the model with monthly data for 2008 – 2010, and comparing to the results.

The model has been run for 3 boreholes using monthly data for 2008 – 2010. No significant difference in the model outcome was experienced and therefore no further actions taken.

3. Calibration and validation periods consist of 1 year each (except at PACE). It is unclear how the model is calibrated and validated (except for the n-factors, where the calibration, validation process is described). Give more detail of the calibration procedure/methods used, and the time period for calibration/validation. The calibration time period must be independent from the validation period. It is unclear whether the two periods were separated completely, and thus difficult to trust the validation results. Note that the time period of the data availability is very short for such large extrapolations in time, please discuss that further.

See Part I, Nr. 4

Minor Comments

Minor comments from the Supplement were taken into account and errors corrected in the text without further commenting.

SECTION 2

Subsection 2.3.1 is not necessary, since it treats both permafrost and seasonal frost conditions at all location (i.e. put the content of 2.3.1 in 2.3, such that 2.3 has NO subsections). Subsection 2.3.2 does not belong to Section 2.3 (permafrost conditions), it can

rather be included in Section 3.2 (Model initialization and boundary conditions) or 3.4 (Historic and future temperature data). Subsection 2.3.4 could be included in Section 2.2 (Climate conditions).

Subsection 2.3.1 was included in Subsection 2.3.
Subsection 2.3.2. was moved to Subsection 3.2 where it is discussed in relation to boundary conditions and model initialization.
Subsection 2.3.3 on "Inter-annual Variation" was included in Section 2.3. on "Permafrost and seasonal frost conditions"

SECTION 4

Section 4.3 (Future permafrost development) could also contain two subsections 4.3.1 (Ground temperature) and 4.3.2 (Active layer thickness). Thereby the structure of Section 4 would be more consistent (similar to 4.2).

The two subsections have been inserted to keep the same structure as in Section 4.2.

SECTION 5

We would suggest a section 5.1 (Model uncertainties), and then some subsections 5.1.1 (Snow cover and water content variability), 5.1.2 (Climate change scenarios). Include further uncertainty chapters (see major revisions).

Additional chapters on uncertainties introduced by the generation of historic air temperature data and the simplicity of the model. However, for reasons of simplicity we avoided the introduction of further sub-subsections.

SUPPLEMENT

Page 815 / Line 21: None of the boreholes were drilled on slopes, therefore real differences in aspect on Jetta are not given.

Page 816 / Line 14: Details on Jetta were added to the chapter.

Page 816 / Line 28: see comment on Fig. 2b

Page 817 / Line 19: Paragraph and numbers are clearified.

Page 817 / Line 21-22: Ground temperatures are close to 0 °C throughout the whole borehole. Concerning the signs of degradation, in addition to these warm temperatures an ALT of 10-11m was recorded during S1 and S2 with a tendency to a further increase. More information was added to the text.

Page 818 / Line 10: Sentence and numbers have been claified.

Page 818 / Line 10: Equation corrected.

Page 820 / Line 1: More details on n-factors were added. See comment Referee #2 (Subsection 3.2.)

Page 821 / Line 5 (Subsection 3.2): The paragraph on model initialisation has been rewritten to make the procedure more clear.

Page 821 / Line 12: The linear correlation is better than r2 = 0.9 at all boreholes. Thus the r^2-value given in that paragraph is the minimum from all boreholes. Since the time series from Fokstugu only differs from the borehole air temperatures by an offset due to the lower elevation, the future time series for each borehole are created by linear correlation.

Page 826 / Line 17: Sentence was revised, now stating: *"…permafrost at the summit reaches the current state of permafrost with an ALT of more than 11 m"*

Page 83 / Line 20: We agree that this is not a direct conclusion of our study. It was partly moved to Section "5.6. Implications of the modelling…"

Fig 2b: The monthly deviations to the climate normal were calculated for the PACE borehole. This was added to the figure caption. Further, the x-axis was changed to SEPT-AUG so that the two seasons S1 and S2 are better displayed.

Fig 3 (now Fig. 4): r-values included in figure

Fig 7 (now Fig. 6): Scatter plots in the right panels include data from the validation period only. Ground temperatures from all depths were included in the scatter plots. Details were added to the figure caption.

Fig 10 (now Fig. 9): Figure caption was changed. Certainly permafrost does not thaw completely, but degradation took place so that no more refreezing occurs.

Reply to REFEREE #2

SECTION 1

p 814 line 8-14: This paragraph is too detailed for an introduction. Just focus on synthesizing the aim of the work and explain the method in following section

Paragraph has been shortened and reorganized.

SECTION 2

This section (sub-sections 2.1 and 2.2) must be re-organized to make it more understandable: there are too many repetitions. It should be probably more clear if you start with a description of borehole sites (merge page 814 line 15-19 and subsection 2.1) and then explain which data are collected at each site and time series length (merge page 814 line 19-27 with the first paragraph of subsection 2.2). Remove from subsection 2.2 the first paragraph.

Recommendations to make this section more clear and understandable were integrated in the text. The section has been reorganized and partly rewritten to avoid repetitions.

Page 814 line 23-25: which is iButton sensors distance on the pole? Say something on this topic so that the accuracy of the snow depth estimation can be evaluated

The following paragraph was added/modified in the text concerning iBotton placement and snow cover extraction:

"In addition to GTs, GST, T_{AIR} and snow depth (SD) is recorded at most boreholes. Low-cost iButton temperature loggers (± 0.5 °C accuracy) were installed at fixed heights (10, 20, 30, 40, 50, 60, 80, 100, 120 cm above the ground surface) along a pole of the air temperature sensor. Using the daily temperature variation, these measurements can be used to determine if an iBotton sensor is covered by snow (Lewkowicz, 2008). This enables the estimation of snow depth development with the accuracy mainly being determined by the vertical spacing of the sensors."

Subsection 2.3.2 this section does not describe permafrost conditions at your sites: consider the idea to move it in section 3.2 where you introduce the concept of GST as upper boundary condition given also the fact that in figure 3 you show modeled data without having introduced the model yet.

This chapter has been integrated in chapter 3.2. 'Model Initialization and Boundary Conditions'.

Change the order of figure 3 and 4: in the text you mention figure 4 before figure 3 and thus consider the idea to change their order.

Accordingly order and numbering of Fig. 3 and Fig. 4 have been changed.

Subsection 2.3.3. In this part your are mainly discussing GST variation between the two years and so you can move it in sub-subsection 2.3.1 (that now can simply become 2.3), where you are already discussing ALT and MAGT variations between the S1 and S2.

Recommendations were integrated and the subsection restructured.

SECTION 3

Subsection 3.2.

page 820 line 13-18: Explain more clearly (or at least give some pertinent reference) how you can derive GST time series using Tair and n-factors.

More information on the calculation of n-factors, their importance and the method to create GST series using them was added to the chapter:

"Historical and future time series of GST were generated from the reconstructed T_{AIR} and downscaled climate change scenarios respectively using n-factors. n-factors can be considered as transfer functions between air and ground surface temperatures during frozen (n_F) and thawed (n_T) state. The n-factors were derived from measured time series of GST and T_{AIR} time series at each borehole by calculating the ratios of annual sums of freezing (FDD) and thawing degree days (TDD) of GST to those of T_{AIR}:

$$n_F = \frac{FDD_S}{FDD_A} \qquad (2)$$

$$n_T = \frac{TDD_S}{TDD_A} \qquad (3)$$

where indices S and A refer to the temperature at the surface and of the air, respectively. Sites with a thick snow cover are characterized by a GST > T_{AIR} during large parts of the winter and therefore n_F < 1. n_T > 1 indicates a warmer GST than TAIR during summer, which can be the case at bedrock sites in the absence of vegetation."

page 820 line 18-21: how can you demonstrate that air temperature and snow conditions observed in S1 and S2 represent the variability of the last decades? Find very pertinent references or smooth this sentence

see Part I, Nr. 3

page 820 line 28: is figure 5 necessary? Consider the idea to simply provide fitting accuracy statistics and eliminate figure 5

Figure 5 was removed from the paper and relevant fitting parameters and accuracies were added to the text in Chapter 3.2 "Model initialisation and boundary conditions"

Subsection 3.3.

Calibration and validation results need to be treated separately.

The Subsection 3.3 was divided into a new Subsection 3.2. Model calibration and 3.3 Model validation. Then calibration and validation is treated individually both in terms of data and text.

Considering the manual tuning of ground parameters: which is the accuracy statistic you used to evaluate model agreement when varying parameter values? To which extent ground parameters influence model accuracy?

See Part I, Nr. 4

Considering the validation of the n-factor approach and table 3: to evaluate model accuracy the use of the Nash–Sutcliffe model efficiency coefficient (EF) is preferable. EF is a measure of the coincidence between observed and modelled data and it is sensitive to systematic deviation between model and observation. See Janssen and Heuberger (1995) for more details.

Following the reviewers recommendation, the Nash-Sutcliffe model efficiency coefficient was used to express the model's accuracy instead of r^2 and RMSE. Comparing measured and modelled GTs values of 0.7-0.9 were achieved.

Table 3: present separately the accuracy statistics for calibration and validation. Use the same statistics for GST, GT and ALT rather than showing r2 and RMSE for GST and GT and modelled and observed data for ALT. GT in table 3 is referred to which depth?

Table 3 was modified and the accuracy statistics (Nash-Sutcliffe model efficiency coefficient) for the calibration period was added.
The ALT refers to the maximum depth of seasonal thaw during summer. Having two seasons available for validation this results in only two values. It therefore is not possible to calculate a RMSE or Nash-Sutcliffe efficiency coefficient. Therefore, the maximum depth both modelled and measured are presented in absolute values. Further, since GTs and GST are in a different unit than the ALT, the RMSE could not be compared directly. The GT and its accuracy measures in Table 3 refers to an average of all depths, thus giving an overall estimation of the model accuracy.

Figure 7: Right panel: plotted data are referred to calibration or validation period? Do not present them together. I suggest again to use EF rather than R2. Is model accuracy dependent on depth?

The scatter plots in the right panel use data from the validation period only. It is now explicitly explained in the figure caption. Nash-Sutcliffe efficiency coefficient is now used. The model accuracy slightly increases from the surface to lower depths.

Figure 7 caption: What does "Only one season was available for calibration and validation for Juv-BH4" mean? You have calibrated over the period Sept-Feb and validated over the period Mar-Jul?

Due to a large data gap only one complete year of measured data is available from Juv-BH4. Since it is essential to have a full typical cycle of freezing and thawing, it was not possible to divide into calibration and validation period. Therefore, for BH4 the one year was used to adjust ground parameters and assess the model accuracy. However, Juv-BH4 is a pure bedrock borehole with negligible water content, no sediment cover nor snow cover. Calibration is relatively easy since the density of the rock, water content and its according heat capacity and thermal conductivity are known and hardly vary in space and time. In addition, Jet-BH3 has the same characteristics and could be used for comparison.

SECTION 4

Subsection 4.2.3

ALT increase values: what does cm a-1 means? Does it mean cm per years? If yes maybe it's better to use cm yr-1.

ALT increases are expressed in cm per year. We changed the format to cm yr^{-1}.

In this section many ALT increase trends (cm a-1) are reported. How this trends have been estimated and are they all really significants? Have you checked them with nonparametric Sen slope estimate or Mann–Kendall nonparametric trend test, or other methods different from linear regression?

The following paragraph was added to Chapter 4.2.3. concerning the calculation fo trends:

"Trends of ALT increase have been derived for the two periods 1860/64 – 1995/1999 and 2000-2010. For the longterm period, the trend was calculated by considering the change of 5-year means of ALT at the beginning and end of the period whereas the trend for the subsequent period was simply derived from the the difference between ALT in the year 2000 and 2010. The nonparametric Mann-Kendall test has been used to test these trends for significance."

The trends were tested for significance using the Mann-Kendall test and the following paragraph was added to the same chapter:

"At all boreholes at Juvvass and Jetta trends of active layer increase during both periods have been proven significant at the 1% level. In contrast, no significant trend was found in the historical development of ALT at Tron until the end of the 1980s. However, the trend over the period of 1990 until today is significant at the 1% level."

SECTION 5

Subsection 5.1

You discuss the effect of n-factors on GST estimation (line 14-23). Consider also the idea to evaluate the effect of different n-factors values on GT and ALT estimations.

See Part I, Nr. 3.

Reply to REFEREE #3

1. **Page 817, line 26-27. The authors say about active layer thickness of 10.7 and 11.1m. Is it still freeze up during the winter time? Does the winter frost penetration reach 11 m?**

   At Tron BH1 we have observed an active layer down to a depth of 10.7 and 11.1 m in 2008/2009 and 2009/2010, respectively. Until today the ground has been refreezing after summer and therefore, active permafrost is still present at Tron BH1.

2. **page 821, 3.3 Model calibration and validation. It is too short period for model calibration (2009/2010). Can you provide more details on procedure of model calibration?**

   See Part I, Nr. 4

3. **Page 822, Section 3.4 (Historic and future temperature data). The methodology of the historic air temperatures producing is unclear for me.**

   See Part I, Nr. 1

4. **Page 822, line 16. Why year 1957 chosen like kind of threshold?**

   See Part I, Nr. 2

5. **Page 826, lines 3-4. Authors say about beginning of a talik development and refer to figure 4e, but from the figure 4e I can not see 'talik development'. It looks like seasonal frozen ground.**

   A simple typing and referencing error is responsible for this confusion. On page 826/l. 3-4 the development of the active layer down to a depth of 11m at TRON BH1 is described. Unfortunately, this was referenced to the borehole Jetta BH3 on figure 4, which does show seasonal frost. The error has been corrected and the figure has been rearranged to Fig. 3:

   a) Juv-BH1
   b) Juv-BH6
   c) Tro-BH1
   d) Tro-BH2
   e) Jet-BH1
   f) Jet-BH3

References

Etzelmüller, E., Schuler, T. V., Isaksen, K., Christiansen, H. H., Farbrot, H., and Benestad, R.: Modelling past and future permafrost conditions in Svalbard, The Cryosphere, 5, 1-13, 2011.

Farbrot, H., and Hanssen-Bauer, I.: A simple station-based empirical model for local snow conditions, Norwegian Meteorological Institute, Oslomet.no report 3/2009, 19, 2009.

Farbrot, H., Hipp, T., Etzelmüller, E., Isaksen, K., Ødegård, R. S., Schuler, T. V., and Humlum, O.: Air and ground temperature variations observed along elevation and continentality gradients in Southern Norway, Permafrost and Periglacial Processes, submitted, 2011.

Hanssen-Bauer, I., and Nordli, P. Ø.: Annual and seasonal temperature variations in Norway 1876-1997DNMI, Report . Norwegian Meteorological Institute, 1998.

Hanssen-Bauer, I.: Regional temperature and precipitaion series for Norway: Analyses of time-series updated to 2004, Norwegian Meteorological Institute, 2005.

Isaksen, K., Ødegaard, R., Etzelmüller, B., Hilbich, C., Hauck, C., Farbrot, H., Eiken, T., Hygen, H. O., and Hipp, T.: Degrading mountain permafrost in southern Norway: spatial and temporal variability of mean ground temperatures, 1999-2009, Permafrost and Periglacial Processes, in press, 2011.

Lewkowicz, A. G.: Evaluation of Miniature Temperature-loggers to Monitor Snowpack Evolution at Mountain Permafrost Sites, Northwestern Canada, Permafrost and Periglacial Processes, 323–331, 2008.

Luetschg, M., Lehning, M., and Haeberli, W.: A sensitivity study of factors influencing warm/thin permafrost in the Swiss Alps, Journal of Glaciology, 54, 696-704, 10.3189/002214308786570881, 2008.

Mohr, M.: Comparison of Versions 1.1 and 1.0 of Gridded Temperature and Precipitation Data for Norway, Norwegian Meteorological Institute, 2009.

Nash, J. E., and Sutcliffe, J. V.: River flow forecasting through conceptual models part I - A discussion of principles, J Hydrol, 10 282-290, 1970.

Stieglitz, M., Romanovsky, V. E., and Osterkamp, T. E.: The role of snow cover in the warming of arctic permafrost, Geophysical Research Letters, 30, 1721, doi:10.1029/2003GL017337, 2003.

Tveito, O. E., and Førland, E. J.: Mapping temperatures in Norway applying terrain information, geostatistics and GIS, Norsk Geografisk Tidsskrift - Norwegian Journal of Geography, 53, 202-212, 1999.

Weismüller, J., Wollschläger, U., Boike, J., and Roth, K.: Modeling the thermal dynamics of the active layer at two contrasting permafrost sites, The Cryosphere Discuss., http://www.the-cryosphere-discuss.net/5/229/2011/tcd-5-229-2011.pdf, 5, 229-270, 2011.

Williams, P. J., and Smith, W.: The Frozen Earth, Cambridge University Press, 1989.

Zhang, T., Barry, R. G., and Haeberli, W.: Numerical simulations of the influence of the seasonal snow cover on the occurrence of permafrost at high latitudes, Norsk Geografisk Tidsskrift - Norwegian Journal of Geography, 55, 261 - 266, 2001.

Zhang, T.: Influence of the seasonal snow cover on the ground thermal regime: An overview, Rev. Geophys., 43, doi:10.1029/2004RG000157. , 2005.