Journal: Hydrology and Earth System Sciences
Title: Projected decrease in wintertime bearing capacity on different forest and soil types in Finland under a warming climate
Authors: I. Lehtonen, A. Venäläinen, M. Kämäräinen, A. Asikainen, J. Laitila, P. Anttila and H. Peltola
MS No.: hess-2017-727
MS Type: Research Article
Iteration: Second review
Referee #1

We would like to thank this referee for the positive feedback. Our replies to the comments are given in "Italics" after each specific comment.

The manuscript has improved significantly since the first review and I recommend publication after some technical corrections:

1) The language can benefit from the help of a native speaker. The manuscript is well written and easily readable but here and there some improvements in the English language will benefit the presentation and overall impression.

Copy-editing service is provided in this journal as a part of the publication process.

2) Could consider to include a reference to damage to the forest terrain in the first sentence of the abstract, and to climate in the last sentence on page 1.

We have added into the end of the abstract a mention that "This is also needed to avoid unnecessary harvesting damages, like rut formation on soils and damage to tree roots and stems."

3) I agree with the approach in Section 2.1.3 wrt subzero freezing point but it is not entirely clear how this affects the comparison between observed and modeled days with frozen soil, cf. the title of the section - Validity of the modeled soil temperatures.

We have added a mention that this assumption concerning subzero freezing points was used in model calculations only. From the observations, it was not so obvious either to estimate the number of days with frozen soil because from most of stations observations were available only every fifth day and the observed soil temperature was often 0.0 °C in winter. These cases were considered non-frozen. In the model, however, cooling of soil does not stop when temperature reaches 0 °C, although it remarkably slows down due to C_{ICE} but this slowing of cooling does not start before soil temperature has already dropped below 0 °C. This is basically the reason why it is not meaningful to consider the soil being frozen right after the modelled soil temperature has dropped below 0 °C.

Journal: Hydrology and Earth System Sciences
Title: Projected decrease in wintertime bearing capacity on different forest and soil types in Finland under a warming climate
Authors: I. Lehtonen, A. Venäläinen, M. Kämäräinen, A. Asikainen, J. Laitila, P. Anttila and H. Peltola
MS No.: hess-2017-727
MS Type: Research Article
Iteration: First review
Referee #2

We would like to thank this referee for the in-depth review. Our replies to the comments describing the changes we have made into the manuscript are given in "Italics" after each specific comment.

Overall assessment

I believe that the study is interesting and of potential interest to the readers of HESS. However, though the manuscript has improved during the revision process, there are still significant issues that need to be addressed. My main concern is the poor description of the methodology, which is of particular interest to the audience of this journal. I find the description of the methodology quite confusing and lacking important information as described in more detail below. Some aspects of the results and discussion are also not clearly explained (details also below), which is at times aggravated by poor English usage.

We have improved the description of the methodology with the help of specific comments provided by the reviewer.

Detailed comments linked to text

The authors refer to "model optimization" throughout the manuscript (starting with the abstract). They are indeed optimizing parameters but with the goal of "calibrating" the model. So this should be clear in the abstract and throughout the manuscript, because within our community model optimization often refers to a very different aim (optimizing water allocation, etc); while here the use of optimization is very clearly focussed on calibration.

We have changed the vocabulary as suggested.

Sections 2.1, 2.2.2, etc.: titles for these subsections: It would be clearer to use the term "description" rather than "outline".

We have replaced "outlines" with "description" in the titles.

Section 2.1.2: Unfortunately, the description in this section is vague and unclear.

Page 5, Line 9, include the soil types for "Lompolojänkkä and Kaamanen ".

We have clarified that "fens" are minerotrophic peatlands.

Page 5, line 10-12: “measurements needed in the calculations were not available from Lettosuo, Lompolojänkkä and Kaamanen stations”, the next lines mention that nearby stations are used, but it is not clear what stations.

We have clarified which stations were used.

Page 5, Lines 13-30: “the optimal parameter values were set to each station and each available measurement depth.” This sentence is unclear. Do you mean that you calibrated the parameters for each station at different depths?

Yes. We state now that “the parameter values were calibrated for each station and at different measurement depths”

What parameters were “calibrated”? It seems that you include the parameters used in the calibration in Table 2, but you should mention them in the text, with a justification of why are these parameters selected (are you calibrating all possible parameters?).

Table 2 includes the calibrated parameters. They are listed also in the text on page 4, lines 24-25.

You also need to justify the selected parameters ranges for calibration.

The sampling ranges were adopted from Jungqvist et al. (2014) but for K_T the upper limit was extended from 1 W m^{-1} K^{-1} to 2 W m^{-1} K^{-1} to better represent the range of soil types and measurement depths considered in our study (page 5, lines x-x)

Are you using a Monte Carlo approach?

We assume that the approach can be described as “Monte Carlo approach” as according to our knowledge, “Monte Carlo approach” refer to almost any approach using random guesses in searching the solution. In principle, the calibration approach was adopted from Jungqvist et al. (2014) where they describe the approach as “a Monte Carlo sampling technique”, although they use different terminology as the procedure was for them “model optimization”, not “calibration.”

Please include a description of the calibration approach, with more technical details, with a more rigorous technical language as this will be easier to follow. For example :

lines 10-20 state: “For some parameters optimized values varied rather randomly within the sampling range between different depths and locations while, for example, KT seemed to steadily increase with soil depth, assumedly largely due to increasing soil moisture with increasing depth”

The calibration procedure needs to be described with more rigour, it is unclear what the behaviour of each parameter was... “some parameters varied randomly” which parameters? Why?

Line 21: “Zl was set to 6.8 m and fS to 9.0 at each location and depth” why these values?

6.8 m was the average for calibrated Z_l values over all the stations and all the measurement depths. In practice, the effect of heat flow from Z_l was negligible at the depths near the surface. Considering f_S, we noted that the calibrated values varied between 9 and 10 with soil depths below 50 cm except at two stations. With increasing soil depth, calibrated f_S values tended to decrease, implying that the effect of snow cover in controlling soil temperature decreases with increasing soil depth, as expected because

the relative importance of heat flow from the surface compared to the heat flow from Z_l decreases with increasing soil depth.

Line 23: "after this second optimization round, all other parameters except KT were also set to their final values." What do you mean by "final values"?

We mean that these parameters were kept constant afterwards.

Lines 23-24: "KT,LOW and CS,LOW were given the same values at all depths and locations while CICE was set to depend on the soil type and CS the depth following asymmetrical sigmoid function." There is no rationale on why this was done.

We have added explanations for these choices in the text.

Line 25: It seems that in the end only Ks is calibrated?

Only K_T was sampled in the final phase while other parameters were kept constant. This was done because less important parameters could have very variable calibrated values at different stations and depths but in order to describe different soil types with fixed parameter values, we needed to have fixed values for each parameter. As K_T was apparently the most sensitive parameter – i.e., whatever values were given to the rest of parameters, a relatively good model fit could be always achieved by tuning K_T values appropriately – in the final phase we sampled only K_T. We have extended discussion related to this choice. We have moreover added to the Discussion section note that a model with almost equally good fit could be achieved with many different set of parameters "because there is no single best model parameter set but many model state descriptions can generate equally good calibration outputs (Beven, 2006; Jungqvist et al., 2014)."

Line 26: "Anjala, Sodankylä and Lettosuo stations were selected to represent clay/silt, sand and peat soil types, respectively" At what point in the calibration was this decided? Why? Shouldn't this be decide at the beginning?

This was decided mainly based on the results seen in Fig. 1, but before fitting the regression curves for the calibrated K_T values. Basically, all the stations were used in the calibration before that phase but at these stations, the calibrated K_T values at different depths followed nicely the logistic regression. Moreover, at some of the stations there were variability in the soil type with different depths so they were not equally well representative for a specific soil type (this was a possible reason for the less clear relationship between soil depth and calibrated K_T values at these stations) and additionally, there were no missing measurement depths at these stations. The explanation has been added to the text.

Lines 27-34: The rest of this section is also unclear.

We have regrouped the rest of section. In the revised version, we only explain how soil frost conditions on forest truck roads were calculated.

Page 6, Section 2.1.3. Please clarify the procedure used here. The authors state:

Lines 2-5: "Apart from the stations used in calibration of KT for different soil types (Lettosuo, Anjala and Sodankylä), the modelled soil temperatures for clay/silt and sand soil types typically explained 90–99% of the observed variability in soil temperatures between the depths of 20 and 100 cm (Table S2)."

This sentence is unclear. You mention “apart from the stations” this suggests that you are going to use data from “other stations”?

Yes, this refers to all other stations in the Table S2 than those three stations mentioned in the brackets. We have rephrased the sentence to make this clearer for the readers.

Lines 5-6: “In spite of the generally high correlations, the modelled number of days with soil temperatures below 0 °C were still greatly overestimated, even by more than twofold on many stations.”

Where are you showing this? Please refer to the table/figure explicitly.

This is not shown with numbers in any table or figure. We have clarified that this is a not shown result.

Page 7. Line 14. You are probably referring to the next section not chapter.

We are not either aware where this sentence is referring to, so we have removed the whole sentence. A reviewer asked us to include this sentence during the previous peer-review round.

Page 10. Lines 11-14 states: “The calculations for the period 1980–2099 under changing climate were completed using daily data from six GCMs (listed in Table S4) participating in the CMIP5 (Taylor et al., 2012; Flato et al., 2013), as bias-corrected and downscaled onto the Finnish grid”. Please clarify: Table S4 does not contain much information. What are the variables/data that are being used? What is the “bias-correction” (note that you mention this later in some figures without any explanation. What is the downscaling? Please add a brief explanation, and if needed more info.

As most of these issues were discussed in the last paragraph of this section, we have reorganized the section and extended this discussion. We have moreover added a reference to Maraun and Widmann (2018) for those readers who would like to have a more in-depth look into the topic of downscaling and bias-correction.

Page 11. Section 3.1: Figure 2 is discussed in section 3.1. It presents the results of the model on number of days with satisfactory bearing capacity for 1981-2010, which are compared to observations. Lines 14-19 state: “the used models generally reproduce the spatial pattern of wintertime bearing season length during the baseline period as expected as the model data has been bias-correct”. I am not sure what patterns are discussed here. The comparison of results from GCM/RCM to those using observation do not show a “general good agreement”. I would argue that the agreement is very poor. The patterns based on observation show a gradual trend from southern to northern portions that is absent from the ones produced by the ensemble forecasts, these last ones have a more random pattern. Please clarify. Please also explain the potential impact of this lack of agreement on the projected trends for climate change.

We have slightly modified the word choices in this section. However, we argue that the agreement between multi-model ensembles and observations is generally good as the number of days with good bearing capacity varies mainly between 60 and 210 days and the differences are in the case of GCM ensemble almost everywhere less than 5 days (with the except of pine forests on peatland) and also in the RCM ensemble only locally more than 10 days.

Page 11. Section 3.2: Lines 25-26 is unclear. It is also not explained why the focus has already shifted to drained peatlands as opposed to the other soils. The discussion should focus on all soil types for this figure.

We have modified the discussion related to these figures by removing the last sentence from the first paragraph because the differences in projected shortening of bearing season length are still small between the different forest types during the near-future period 2021–2050.