Response to comments of reviewer #1:

Major Comments:

1. This study used Wang06 permafrost map as a reference. As compared with Wang06map, there is a tendency that the indirect methods (MAAT, F, SFI) overestimated the permafrost area (Table 3). For the calculation, authors used thresholds of -2^{o}C MAAT and frost number ≥0.5. It seems their selection was considerably subjective, based on empirical results of previous studies. **The thresholds have considerable potentials that can change the results of this study.** For example, when MAAT is set to 0^{o}C, the permafrost area becomes probably larger than the areas indicated on Table 3. **That is, the results of this study also include latent uncertainties in terms of the methodologies.** A possible way to reduce the uncertainty is what examines the sensitivity of permafrost area against the changed threshold, calculating changes in the permafrost area to the ranges from -3 to 1^{o}C of MAAT. The calculations have to be made for frost number. The have to be summarized as a table and figure, including the discussions.

We have done as the reviewer suggested and added extra thresholds as a new column in Table 2 (i.e., -3 ℃<MAAT<0 ℃; 0.4<F<0.6; 0.4<SFI<0.6; 0 ℃<MAGT<0.5 ℃). And the derived permafrost area is now listed as extra rows in Table 3.

Generally, when the permafrost definition requires colder climate, the derived permafrost area becomes smaller. The across-threshold uncertainty (Table 3) is similar for different models. But the across-threshold uncertainty with SFI varies greatly among models, 23 ~ 105 ×10^{4} km^{2}, which is due to the seasonal amplitude of ground surface temperatures it requires. This is illustrated in Fig. 5 where UVic and LPJ-GUESS have a relatively small seasonal amplitude of ground surface temperature, which corresponds to their small across-threshold variability for SFI derived area in Table 3.

The across-model uncertainty is highly consistent even with different thresholds for each method (Table 3 final column). Thus it seems changing the thresholds does not affect one key point in our paper: that across-model uncertainties using direct methods are much larger than using indirect ones. Large across-model uncertainties using direct methods imply that differences among these land surface processes are worthy of investigation.

Accordingly, we add this analysis in the new discussion in section 6.1.

Revised Table 2.The five diagnostic methods and threshold values used to derive permafrost. The thresholds commonly used in the literature and in this paper are marked in bold.

Method	Definition	Threshold	Data used for calculation
TSL	More than 24 consecutive months soil temperature ≤ a threshold	**0 ℃**	0 ~ 3m monthly soil temperature
MAGT	Mean annual of 3 m soil temperature ≤ a threshold	**0 ℃**, 0.5 ℃	Mean annual of 3 m soil temperature
SFI	Surface frost number ≥ a threshold	0.4, **0.5**, 0.6	Annually maximum and minimum ground surface temperature

F	Air frost number ≥ a threshold	0.4, **0.5**, 0.6	Annually maximum and minimum air temperature
MAAT	Mean annual air temperature ≤ a threshold	0 ℃, -1 ℃, **-2 ℃**, -3 ℃	Mean annual of air temperature

Revised Table 3. Derived permafrost area inside the common modeling region on Tibetan plateau (10^4 km^2) from 6 LSMs and 5 diagnostic methods, using different thresholds. The results of thresholds commonly used in the literature and in this paper are marked in bold.

		CLM4.5	CoLM	JULES	UVic	ISBA	LPJ-GUESS	across-model uncertainty
Indirect method	MAAT≤ 0℃	130	124	126	116	127	129	14
	MAAT≤ -1℃	122	117	119	109	119	120	13
	MAAT≤ -2℃	**113**	**105**	**111**	**99**	**109**	**110**	**14**
	MAAT≤ -3℃	95	83	96	81	91	93	15
	across-threshold uncertainty	35	41	30	35	36	36	
	F≥ 0.4	140	135	138	126	138	138	14
	F≥ 0.5	**135**	**127**	**131**	**118**	**130**	**131**	**17**
	F≥ 0.6	117	93	106	89	100	101	28
	across-threshold uncertainty	23	42	32	37	38	37	
Direct method	**TSL**	**60**	**1**	**62**	**8**	**44**	**119**	**118**
	MAGT≤ 0.5℃	112	102	104	8	72	131	123
	MAGT≤ 0℃	**104**	**89**	**96**	**8**	**61**	**128**	**120**
	across-threshold uncertainty	8	13	8	0	11	3	
	SFI≥ 0.4	135	122	130	32	131	127	103
	SFI≥ 0.5	**116**	**62**	**100**	**8**	**113**	**119**	**111**
	SFI≥ 0.6	42	17	38	4	55	104	100
	across-threshold uncertainty	93	105	92	28	76	23	
across-direct method uncertainty (based on commonly used methods TSL, MAGT≤0ºC, SFI≥0.5)		**56**	**88**	**38**	**0**	**69**	**9**	

2.a) Snow cover has the insulation effect on soil temperature; deeper snow depth increases soil temperature and vice versa. However, figure 7 and 8 have results inconsistent with the facts, especially found in ground surface temperatures of UVic, ISBA, and JULES. They are not common. The explanations authors mentioned are not enough to clear up the question.

We agree with the basic snow insulation mechanism, but UVic behaves differently from that of ISBA and JULES.

In the case of UVic, the ground surface temperature was warmer than the air despite of no snow cover.

There are two aspects that account for UVic behaviour: albedo and sublimation. Fig. 6 now more clearly shows that UVic simulates no snow (even in winter). Since there is no snow, the albedo is quite low all year round, 0.15-0.35. This means UVic can absorb much more solar radiation in winter, which can greatly warm the ground.

However, since the forcing of UVic is the same as CLM4.5, there is snowfall. Then where does the snow go? We assume the sublimation removes the snow in UVic. If so, it should take more energy to sublimate snow than it does to melt it, then the latent heat flux should be higher in UVic than other models. Actually we did find that the evaporation rate in UVic is higher.

In conclusion, the low albedo leads to more solar radiation absorption, which warms the ground and provides energy for sublimation. Although sublimation can cool the ground, the warmer ground indicates the low albedo effect is stronger in UVic.

We have addressed this in section 5.

In contrast, the surface temperatures of ISBA and JULES were colder under thicker snow cover.

Yes, observations show that in general deeper snow depth increases soil temperature. However, if the snow is too thin, the warming effect is very weak. But the melting, evaporation, and sublimation processes will take much energy from the ground, and the snow cover will reflect much solar radiation. Both will cool the ground. That is the thin snow cooling effect.

Fig.6 shows that in most places on Tibetan Plateau the snow depth of ISBA and JULES is less than 10 cm. Thus, the colder ground temperature of ISBA and JULES may be due to the thin snow. We investigate this in the new plots Fig. 9, which is like Fig.7 in the paper, but shows the temperature offset between ground surface and air temperature for different snow depths. By inspection we note that there is different behavior for snow depths thinner and thicker than 4 cm. For snow depth > 4 cm, most negative offsets disappear in ISBA and JULES, which means that the ground surface temperature is warmer than air temperature for snow depth larger than 4 cm. For snow depth < 4 cm, the ground surface temperature of much of the region is colder than air temperature in ISBA and JULES, which indicates the cooling effect of thin snow. The very small or slightly negative temperature offset for thin snow is also seen in the other models. Of course, the strength of this effect depends on the individual model's simulation/parameterization of the snow processes (such as sublimation, evaporation, melting). The thin snow mechanism is also confirmed by the weak insulation effect in Fig. 10.

Accordingly, we have improved the discussion of these issues in section 5. We also add Fig. 9 and Fig. 10.

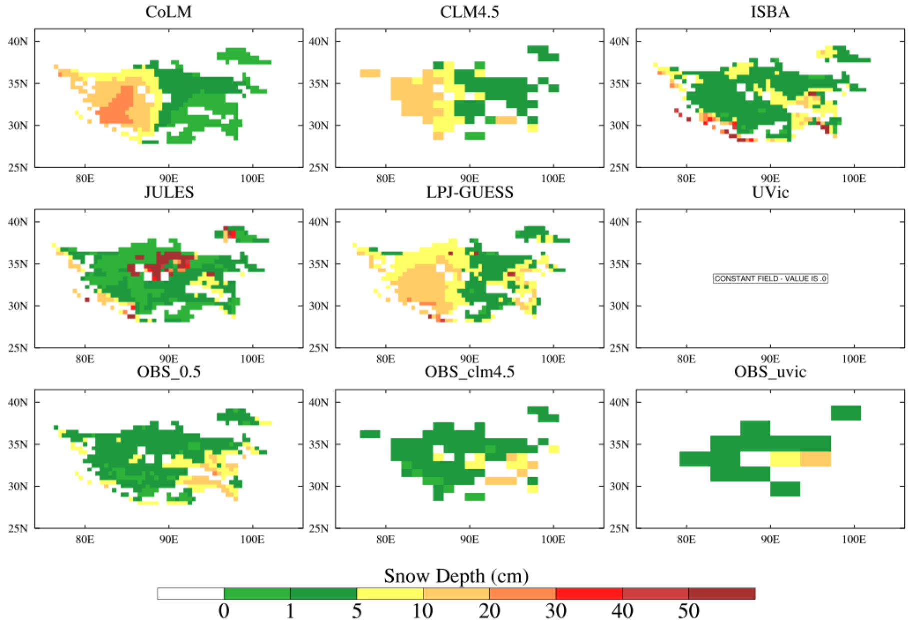


Figure 6. *Winter snow depth for the common region, averaged over 1980-2000. Note the nonlinear color scale. We use the Long Time Series Snow Dataset of China (Che et al., 2008) (http://westdc.westgis.ac.cn) as observed snow depth. The observed snow depth plot is further interpolated onto the models' resolutions as "OBS_". The OBS_05 is in 0.5 °resolution for CoLM, ISBA, JULES and LPJ-GUESS. The OBS_CLM4.5 and OBS_UVic are in the resolutions of CLM4.5 and UVic separately.*

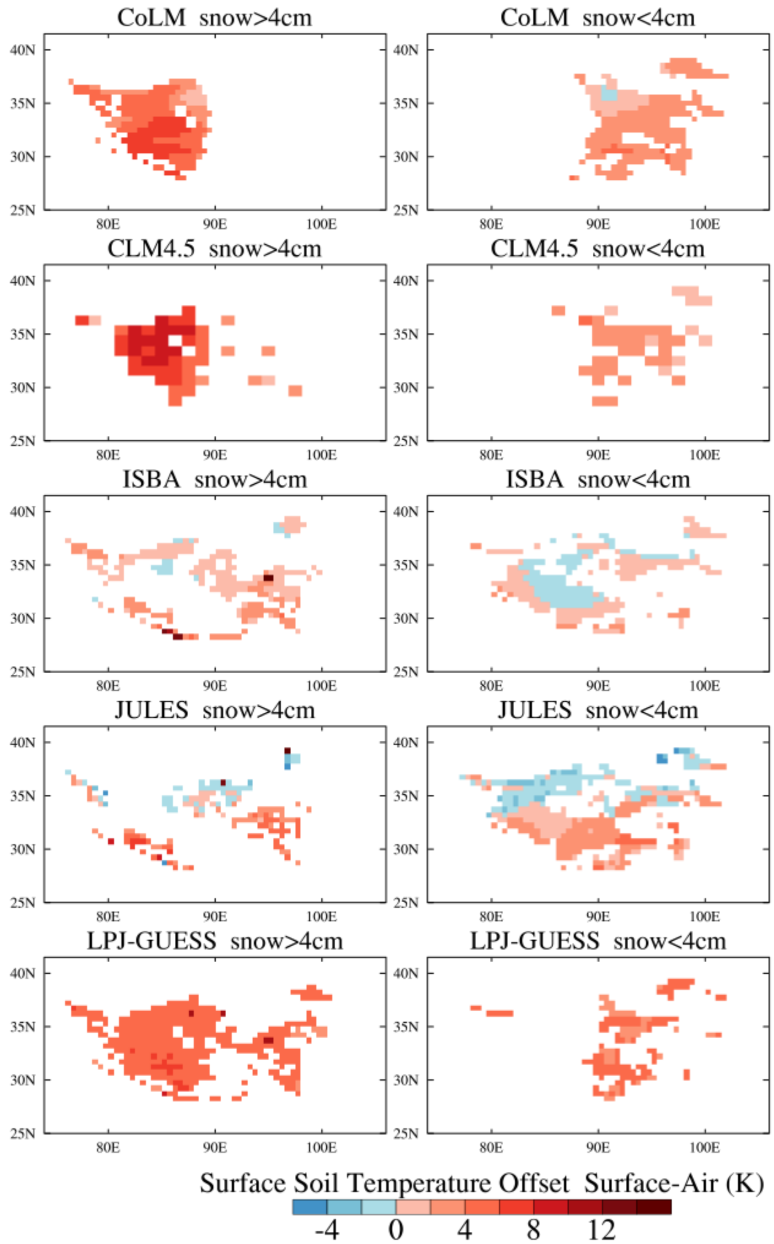


Figure 9. *Mean surface temperature offset (difference in mean winter temperatures between surface soil and air, averaged over 1980-2000). Left column is for snow depth >4cm, right column shows regions with snow depth<4 cm. Warm colors indicate soil is warmer than air temperature.*

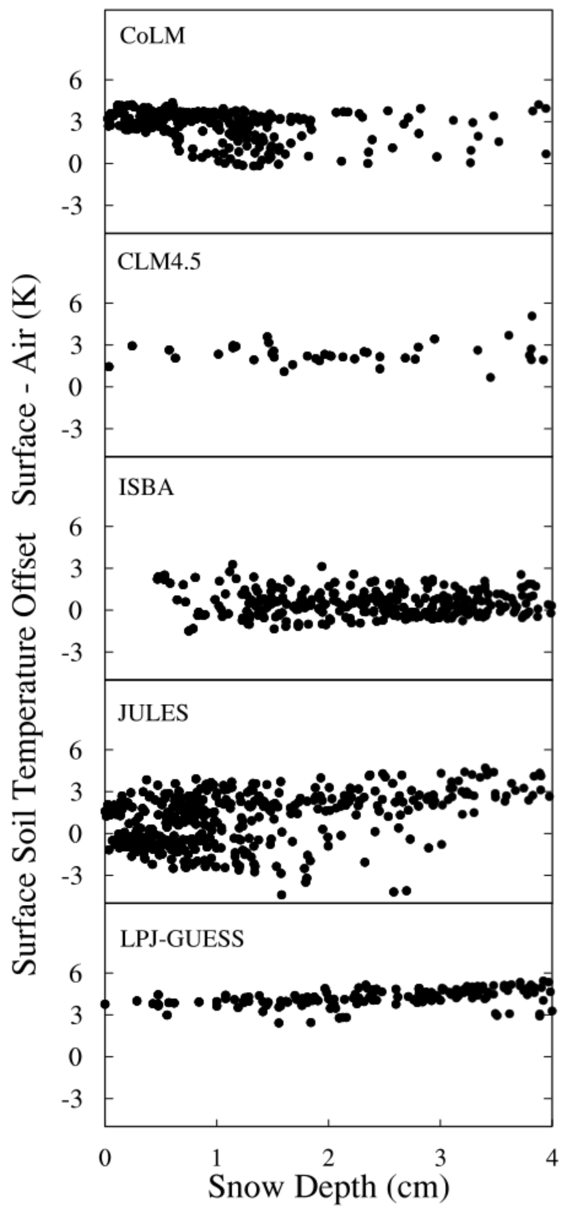


Figure 10. *Mean surface temperature offset (difference in mean winter temperatures between surface soil and air, averaged over 1980-2000) as a function of snow depth for grid points where average snow depth <4 cm.*

For the reasons, the authors mentioned problems of parameters the model used and the reliance of snow cover data derived from satellite observations. Of course, we can enough consider their influences. However, if there is problem in the snow cover data, how do you explain the results of the remaining three models that satisfy the common facts?

According to Wang et al. (2013), the snow depth pattern and the significant seasonal snow characteristics of satellite data are consistent with those of station data in most of our common TP region. The satellite data are different from station data on the southeast of TP (Wang et al., 2013), however, our analyzed common region does not include this part of TP. We add discussion in the second paragraph of Sect.5 where we introduce satellite data.

2. b) The ground surface temperature was extrapolated from the vertical soil temperature profiles. How did you extrapolate it, by liner or exponential way?

Ground temperatures were linearly interpolated onto the common depths: 0.05, 0.1, 0.2, 0.5, 1, 2, 3m. And we use the top two layers' soil temperature to linearly extrapolate the ground surface temperature. For CLM4.5, CoLM, ISBA and LPJ-GUESS, the first layer soil depth is no deeper than 0.01m and the second layer soil depth is no deeper than 0.05m. For JULES and UVic, the first layer soil depth is 0.05m and the second layer soil depth is no deeper than 0.18m. We do not expect that the presented results are sensitive to the interpolation procedure.

Accordingly, we improve these sentences in P1779 L10-2 as "... ...Ground temperatures were linearly interpolated onto the common depths... ...Since there is no ground surface temperature output, we linearly extrapolate the top two layers' soil temperatures onto the ground surface. For CLM4.5, CoLM, ISBA and LPJ-GUESS, the first layer soil depth is no deeper than 0.01m and the second layer soil depth is no deeper than 0.05m. For JULES and UVic, the first layer soil depth is 0.05m and the second layer soil depth is no deeper than 0.18m. Most TP permafrost work... ..."

2. c. Soil temperature doesn't generally have linear vertical profile. As the first step, you have to check the soundness for the extrapolated surface temperature; a way is to calculate the temperature offset using soil temperatures of 0.05 m. If there are no large differences, we could be doubtful about the snow cover data. If not so, you could have to recalculate the surface temperature that is implicated to SFI.

We agree. We plotted the surface temperature offset and soil temperature offset using the soil temperature of 0.05m depth, which can be compared with our Fig.7 and Fig.8, see below. The results show that first, the temperature offset patterns using 0.05m soil temperature are consistent with Fig.7 and Fig.8. Second, the offset between 0.05m soil and air temperature is only slightly different from Fig.7. Third, there is still negative surface offset in ISBA and JULES using 0.05m soil temperature as in Fig.7. This indicates that our interpolation for ground surface data is reasonable.

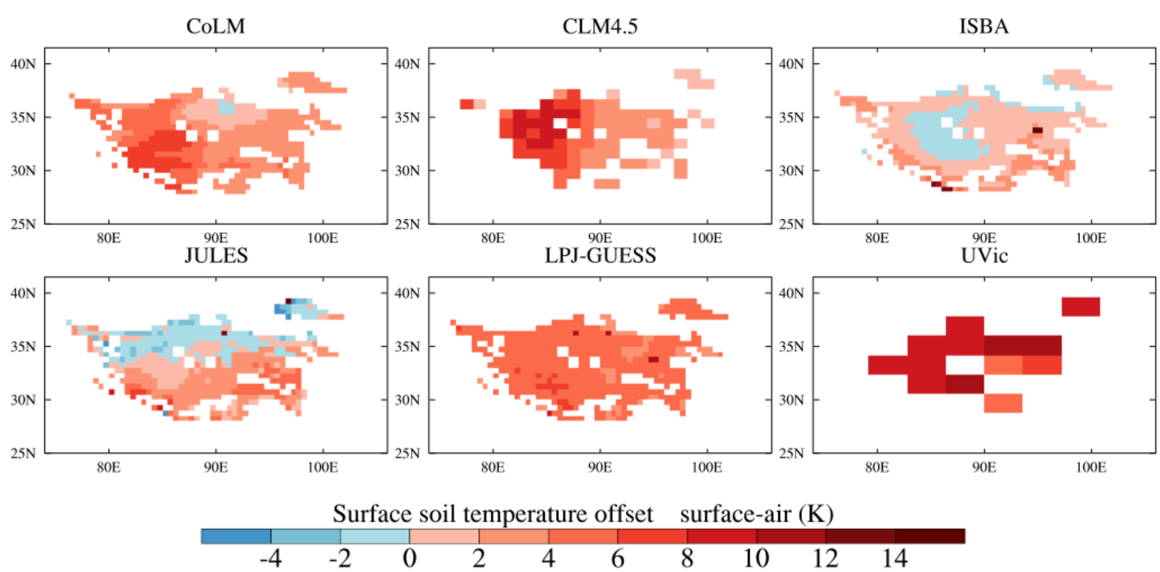
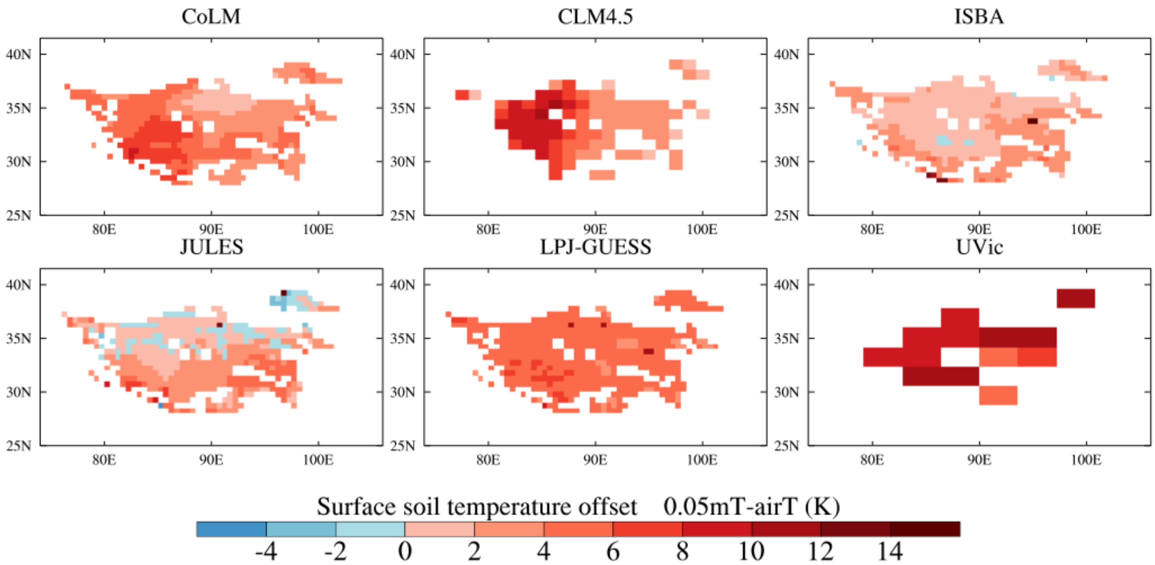


Figure 7. Mean surface temperature offset: difference in mean winter temperatures between surface soil and air, averaged over 1980-2000. Warm colors indicate soil is warmer than air temperature.



*Mean surface temperature offset: difference in mean winter temperatures between **0.05m soil and air**, averaged over 1980-2000. Warm colors indicate soil is warmer than air temperature.*

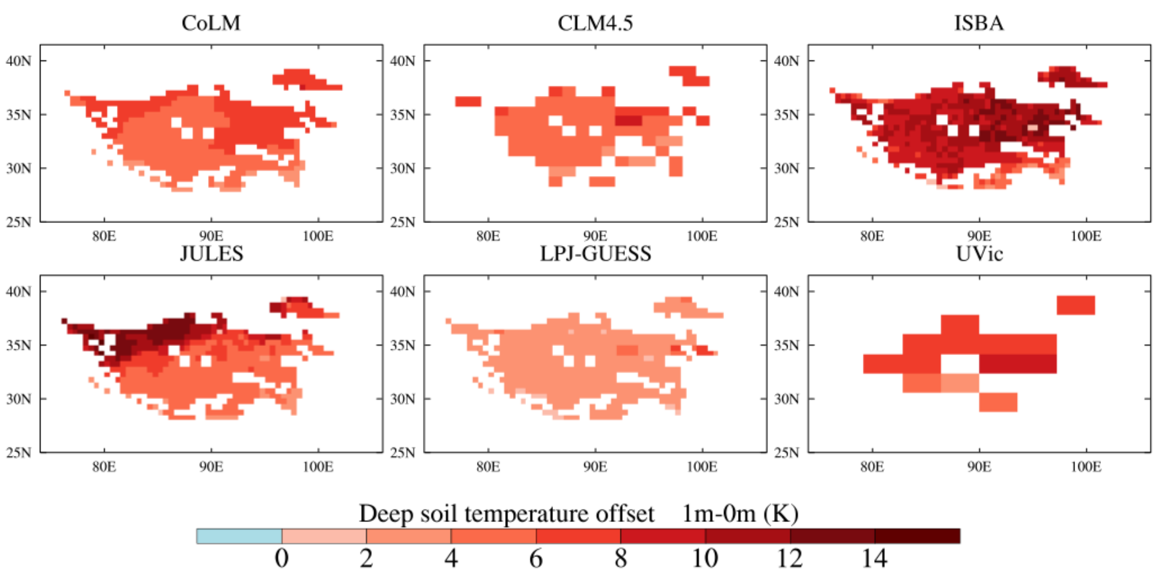
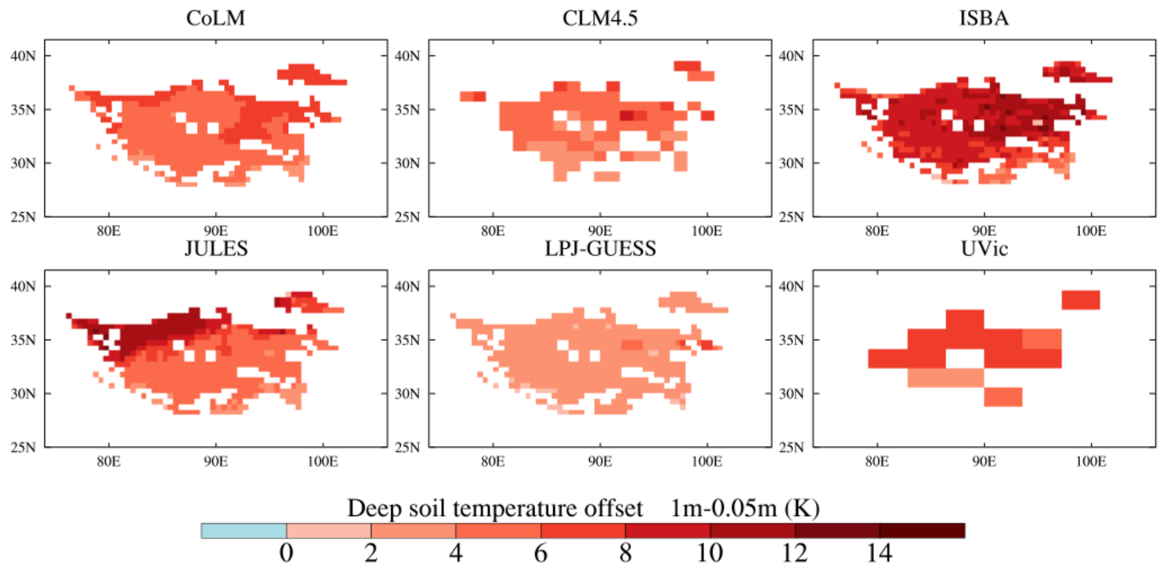


Figure 8.*Mean soil temperature offset: difference in mean winter temperatures between soil at 1 m depth and surface soil, averaged over 1980-2000. Warm colors indicate deep soil is warmer than shallow soil.*



Mean soil temperature offset: difference in mean winter temperatures between soil at 1 m depth and ***0.05m depth soil****, averaged over 1980-2000. Warm colors indicate deep soil is warmer than shallow soil.2.*

d) The authors mentioned influences of model parameters on the problem, but they didn't provide any scientific bases. For example, CLM4 (Lawrence et al., 2012, J. Clim.) and JULES (Dankers et al., 2011, Cryosphere) have characteristics of cold bias for soil temperatures. Likewise, you have to approach for the problem with scientific results and physical properties of individual models.

In this part we understand the referee means the issue of why UVic seemed to give large warming without snow, and also why ISBA and JULES give negative insulation for ground despite snow cover. We explained this in reply to part 2a above: it is due to albedo, sublimation and thin snow cooling effects. This is now explained in section 5, and it seems to us quite similar in method (that is using "scientific results and physical properties of individual models") as suggested by the referee and in the papers by Lawrence et al. and Dankers et al.

3. a) As the authors had already mentioned, the lowest soil boundary is a critical issue implicating to the permafrost uncertainty. Three models among the six models extended the soil to deeper depth, which makes it possible to do the discussion on this issue. The authors described a result of CLM4.5; when the soil was extended to 38 m, there was no significant change in the permafrost area (P1776 L3–5). From this description, the reviewer understands that there was a difference, though the difference was not so large numerically. If so, the authors have to provide the analyzed numbers on the manuscript, including results of other two models. The related description is summarized with one paragraph in the section of Conclusion, but which has to move to the Discussion section with additional discussion.

Agreed, we moved the paragraph to the discussion section and now provide the details of the calculations. As the reviewer suggested, we calculated the permafrost area with deeper soil layers. Note that it is the same simulation for the results of both 3 m and 30 m. " As UVic does not do a reasonable simulation of snow cover and ground temperature, we feel it is not necessary to include this model in the discussion here. Based on results from CLM4.5 and ISBA, the permafrost area calculated from MAGT at 3 m and at 10 m only changes by $1x10^4$ km^2. For results from CLM4.5, the areas calculated from MAGT at 20 m and 30 m do not change from the one calculated at 10 m. This is due to MAGT only considering annual mean soil temperature, not the seasonal cycle. This is consistent with the finding that the across-threshold uncertainty for MAGT-derived permafrost area is quite small (Table 3). However, the derived permafrost area with the TSL method improves when soil depth used for calculation is increased from 3 m to 5 m (Table 6). This sensitivity is because TSL requires information on the seasonal cycle of soil temperature. In other words, results of TSL method are sensitive to the active layer dynamics. The permafrost on TP is usually much warmer and has a deeper active layer than found in continuous permafrost of the arctic and boreal region. Hence deeper soil layers would be well suited for TP permafrost simulation. "

Accordingly, we add this analysis in the new discussion section 6.2, along with the new Table 6.

Table 6. Derived permafrost area (10^4 km^2)with deeper soil layers, using method "TSL". The results of thresholds commonly used in the literature and in this paper are marked in bold.

Depth of deepest layer used for calculation	CLM4.5	ISBA
3m	**60**	**44**
5m	85	54

3.b)The monthly time resolution this study used could also increase the uncertainties for the estimated permafrost area, which has also to be mentioned in the Discussion section.

Agreed. The methods used here are defined for monthly data, but standard definitions are for daily data or hourly data. We do not have that output from the models. We added a sentence in the last of section 6.2: "Daily and hourly temperature data may make some differences to the permafrost extent map since the depth of diurnal cycle wave in soil is much less than the deepest model layer."

Minor Comments:

P1771 L9, the word 'best' is not appropriate. Please remove it.

Done. **It reads now** "There is good agreement (99–135 $\times 10^4$ km^2) between the two diagnostic methods based on air temperature which are also consistent with the observation-based estimate of actual permafrost area (101 $\times 10^4$ km^2)."

P1771 L15-7, it seems differences in vegetation types are implicated to.

Done. We add vegetation types. **It reads now** "Model evaluation at field sites highlights specific problems in process simulations likely related to soil texture specification, vegetation types and snow cover."

P1772 L4, climate likely affects the permafrost distribution.

Done. We add climate factor. **It reads now** "The unique geography and plateau climate make the permafrost on TP very different from the Arctic."

L10, does 'the temperature' mean air temperature?

Here we refer to the underlying surface temperature contrast. **It reads now** "The underlying surface temperature contrast between TP and Indian Ocean is an important controlling factor for both the Asian monsoon, and the wider general atmospheric circulation (Xin et al., 2012)."

L19, where are individual station locations?

The station of VIC model here is on the permafrost of Heilongjiang. Others are all on the Tibetan Plateau. We delete VIC model here. **It reads now** "A number of land surface models (LSMs) (e.g., CLM4.0, CoLM, SHAW, Couple Model and FSM) have been applied at individual station locations on the Tibetan Plateau to reproduce soil thermo-hydro dynamics (Li et al., 2009; Wang and Shi, 2007; Xiong et al., 2014; Zhang et al., 2012)."

P1774 L18-9, please add references.

Done. The references have been **added as** "We make use of all five major permafrost diagnostic methods promoted in the literature (Slater and Lawrence, 2013; Guo et al., 2012; Guo and Wang, 2013;Wang et al., 2006; Wang, 2010; Nan et al., 2002;Nan et al., 2012; Saito, 2013; Ran et al., 2012; Wang et al., 2006; Jin et al., 2007;Xu et al., 2001; Nelson and Outcalt, 1987)."

P1778 L15, what means 'SD'?

Sorry, 'SD' means standard deviation. We **correct it as** "rather than defining uncertainty in terms of standard deviation".

P1778 L26–P1779 L6, I don't well understand why Miroc-ESM was used in this study.

This is a misunderstanding. We analyzed Miroc-ESM since Miroc-ESM is also under "the Permafrost Carbon Research Network". But we didn't put its results in the paper, since Miroc-ESM is totally different from the land models. See P1779 L1 "We also analyzed (but do not show here) out put from a coupled earth-system model (Miroc-ESM)".

Since these sentences are rather misleading, we **deleted them** (P1778 L26 - P1779 L6) "We also analyzed (but don't show here) ... ... with the stand-alone LSMs".

Figure 5 certainly includes data of Miroc-ESM, but the data are not closely related to this study. Your intent is to strengthen the high uncertainty in the permafrost processesbetween models, is it?

Sorry, this is a mistake. Yes, the intent is to strengthen the high uncertainty in the permafrost processes between models. We **deleted** Miroc-ESM in Figure 5.

P177**8** L10-2, you have to describe more specifically the way on how the ground surface temperature was extrapolated.

We think here you mean we should improve the extrapolation method description in P177**9** L10-2. Please see our detailed reply to your question 2 b).

P1779 L17-24, it may be helpful if you describe the specific numbers about TP permafrost area reported by previous studies.

Agreed. These numbers can further indicate the permafrost area uncertainty of the observation-based maps. We **added a sentence** in P1779 L24 **as** "...... different studies (Ran et al., 2012). Thus there is a large spread of observation-based TP permafrost area estimates from 110 × 10^4 km^2 (Wang et al., 2006) to 150 × 10^4 km^2 (Shi and Mi, 1988; Li and Cheng, 1996)."

P1781 L20-1, The description about Figure 5 is too simple. You have to add more explanation about data displayed on the figure. From the figure, we can see differences in the simulated soil temperatures between models. The air temperature also shows differences between the models, especially in winter season, though the differences are smaller.

Done. We added a sentence "The air temperature also shows differences between the models, especially in winter season, though the differences are much smaller than differences in soil temperatures."

P1782 L10, does 'observation' mean Wang06 map?

Sorry, this is a mistake. We changed "observation" to "Wang06 map".

L19-21, we can consider the impact of different forcing data, especially air temperature as identified in Figure 5. Therefore, it can't conclude the difference with just different spatial resolution.

Agreed, and deleted the sentence.

P1786 L20-1, for 'poor representation of soil hydrology', I can't find any scientific results from this paper to support the description.

Well, here we think the soil hydrology of LPJ-GUESS plays two parts in underestimating of soil temperature 1) through affecting soil thermal properties, see P1786 L24-25 "This suggests a different (larger) winter soil thermal conductivity probably associated with a high soil porosity and water content"; 2) through hydrological process, e.g. ice melting in summer, which can be confirmed by Fig. 4a and c, see P1787 L2-7 "Precipitation and hydrological processes determine the vertical profile of soil water content which can change the fraction of water and ice retained in different soil layers and influence soil thermal conduction. The energy required to melt the high water (ice) content in the surface soil layers in summer appears to lead underestimated low summer temperatures compared with other models, and a phase lag in summer warming (Fig. 4a and c)." Thus we think "poor representation of soil hydrology **on Tibetan Plateau**" is solid with the two parts of soil hydrology in LPJ-GUESS.

P1788 L14-6, is the description able to apply to all models?

In fact, CLM4.5, CoLM, JULES and ISBA do produce better permafrost maps with MAGT and SFI than TSL. LPJ-GUESS and UVic are two exceptions, showing us no change from TSL to MAGT and SFI. Here the word "generally" means "most models". But if this is misleading, we would change it.

Accordingly, **we changed this sentence with** "Most models in this study produced permafrost maps in better agreement with the Wang06 map using the MAGT and SFI methods rather than with the TSL method."

Table 3, it may be helpful if the uncertainties compared to Wang06 are included.

We find that showing the uncertainties compared to Wang 06 in Fig.2 works better than in this table.

Figure 5, it may be helpful if the seasonal snow depth derived from the data of Figure6 is added.

We add additional discussion and extra figures on snow depth that we think better explains the points. We address this in detail in our reply to your second comment.