tc-2016-115-RC1

Dear Anonymous reviewer,

Hereafter the responses to your comments.
(note: The number of the figures and tables are not coincided with the revised manuscript. It is just the order in the supplement file)

General Comment:
The paper presents a very interesting application of joint AMSR-E and AMSR2 remote sensing data to freeze/thaw (F/T) patterns in the context of permafrost evolution. Whereas the general usefulness of the approach becomes apparent, which should be further studied and explored, the present state of the manuscript contains many methodological flaws (or at least inaccuracies) which from my perspective prohibit that the paper can be published without very serious modifications and improvements. These I will list in the following.
Response:
Thank you for your positive comment on the usefulness of our technique. We have carefully addressed all your concerns as outlined below.

Comment (1):
First of all, the paper suffers from many inaccuracies of the English language, the format (missing blanks between words, readability of the figures, especially the axes and color scales) and the references.
Response:
We have invited a native English speaker to review our manuscript, and we believe the language and its formats have been improved in our revised version.

Comment (2):
The F/T signal from the remote sensing data is compared with air temperature to assess the accuracy. This very probably introduces a strong bias, as not only soil specific factors are neglected, but also, and more importantly, the isolating effect of the snow cover is neglected! As the snow cover is very variable spatially and temporally, its impact on soil freezing will also be non-trivial and cannot be neglected. For the global application (which does not make much sense in my opinion anyway, see below) air temperature data might be justified, as soil temperature is difficult to obtain, but for permafrost areas such as the Qinghai-Tibetan Plateau, this information is available.
Response:
Thank you for your nice suggestion, we quite agree with you that the air temperature might introduce a bias to assess the soil F/T accuracy. But for the technique of passive microwave remote sensing at a spatial resolution of about 25km, the direct validation is always challenging. In the previous studies of Zhao et al., 2011 and Chai et al., 2014, the accuracy of this algorithm has been compared with 4-cm soil temperature measured over the Qinghai-Tibetan Plateau and 0-cm land surface temperature in the region of China. In this study, we focus on evaluating the continuity and its accuracy of AMSR-E

and AMSR2 data records, therefore, only data that are available over a very long time range and space domain are selected.

Changes:

Together with your comment and other reviewer's suggestion, we further used the in-situ measured (ISMN 0-5 cm soil temperature) and modelled soil temperature (GLDAS Noah 0-10 cm) to assess the accuracy of the F/T dataset over the domain of 26°N-90°N (This is changed to match the main permafrost area). The new results found that the satellite derived F/T map has an agreement 82.25% (Spatial agreement, Figure 1 in supplement file) and 85.43% with WMO air temperature in 2010 and 2013 respectively. Please see Section 2.2 for descriptions of these ground based measurements and other external data, added results please see Section 3.2. The overall agreement with ISMN soil temperature is 86.63% for all the stations in the year of 2010 (Table 1 and Figure 2 in supplement file). The results with GLDAS modelled soil temperature indicate an overall agreement of 89.74% (A) and 87.6% (D) in the research domain (Figure 2).

Figure 1 Comparison with WMO air temperature (2010 and 2013)

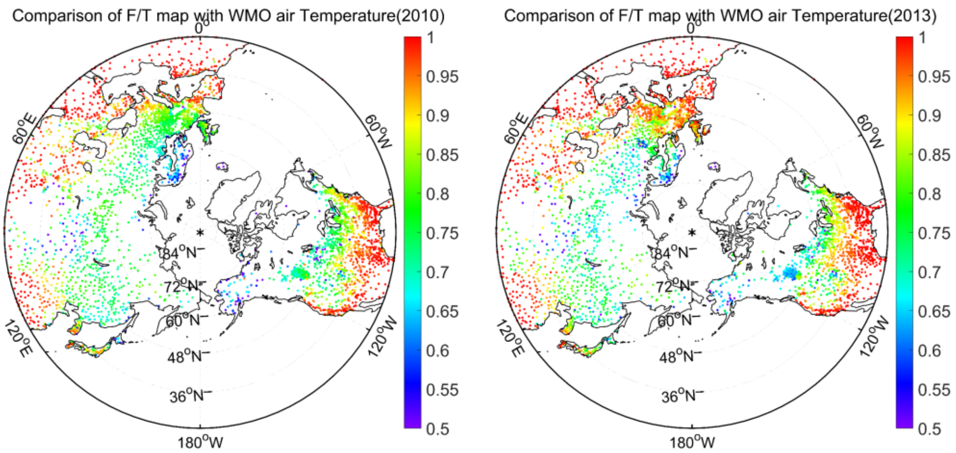


Figure 2. Comparison with soil temperature from ISMN

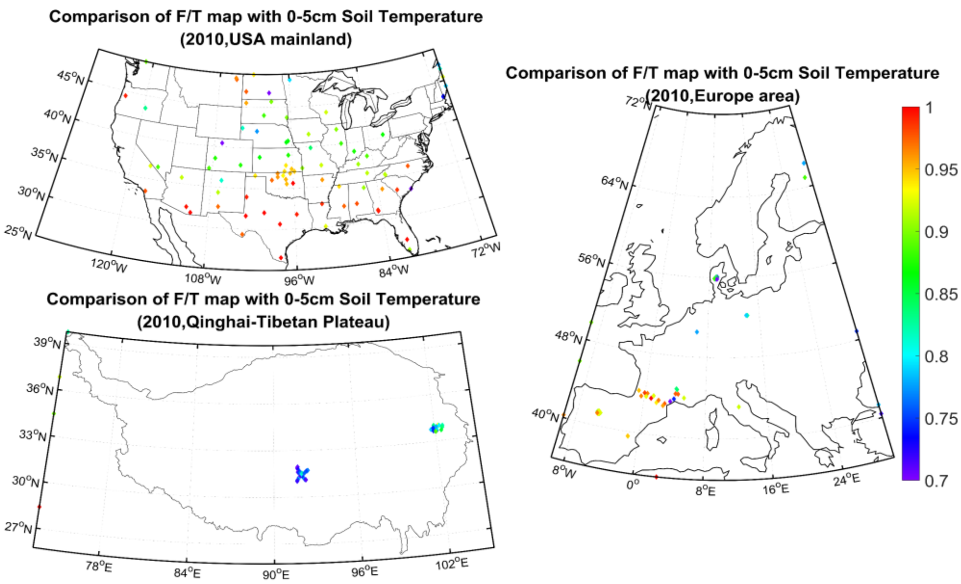
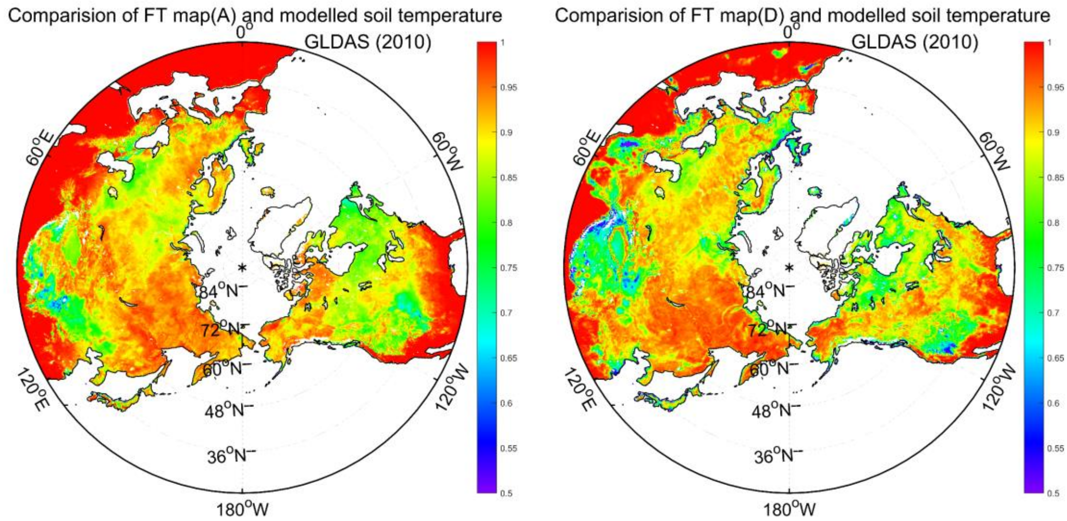


Table 1. The main soil moisture networks (ISMN) distribution and the agreement with F/T map

Main Soil Moisture Networks	*location*	*Total stations*	*Overall agreement*
ARM	USA	88	91.17%
SCAN	USA		
HOBE	Europe	75	87.65%
REMEDHUS	Europe		
SMOSMANIA	Europe		
MAQU	QTP	48	77.20%
CTP_SMTMN	QTP		

Figure 3. Comparison with modelled soil temperature from GLDAS (2010)



<u>Comment (3):</u>

<u>Because of (2), the "accuracy" mentioned is rather a statistical correlation between different variables, as not the same variable is compared.</u>

Response:

Having real "ground truth" data is critical but very challenging. The disagreement between satellite footprint scale (about 25 km in this study) estimates and point scale ground measurements make the direct validation becomes almost impossible. However, all validation activities are assuming the data used are an "accurate" reference to calculate the "accuracy".

Changes:

To make it more reasonable, in our revised manuscript we used "agreement" to replace "validation", but we keep the word of "accuracy".

<u>Comment (4):</u>

<u>The statistical results are given with too many decimals considering the large uncertainties inherent to the approach. Furthermore, many of the statistical analyses are scientifically meaningless without explaining what they are used for (e.g. what does a globally averaged</u>

number of frost days could mean on this level of accuracy and temporal and spatial scale?).

Response:

Understanding the spatial distribution and temporal evaluation of frost events is vital given their impacts on vegetation growth. Ecosystems that do not experience frost such as tropics are highly vulnerable to cold temperature. Alternatively, ecosystems accustomed to long frozen periods are highly vulnerable to warm temperatures. Therefore, it is an important indicator to climate change. The number of frost days calculated in our study is basically for a decade, together with more historical data records, it would be able to explore the changes in permafrost and frozen ground and its relationship with climate warming or ecosystem at a global scale. And because there are uncertainties associated with the approach, we retain decimals as it is a scientific research paper. Similarly reports could be found in many studies (e.g. Kim et al., 2014, Jin et al., 2015).

Changes:

According to your concern, we have added more explanation regarding to frost days (see Section 2.3.4).

Comment: (5):

Page 1, line 29-30: Significant changes in frost days are obtained mainly for regions of discontinuous and transient permafrost: this can potentially be influenced by a strong bias through a variable and changing snow cover, as the latter tend to be changing strongly in these climatic zones.

Response:

We fully agree with you that the changing snow cover is a critical factor impacting the detection of soil freeze/thaw state. According to the criteria selected in our algorithm, the snow-covered ground presents a very similar characteristic with frozen ground, as the difference in volume scattering between 18.7 and 36.5 GHz would cause a very high Quasi-emissivity (may be larger than 1 for very dense snow layer). Therefore, if the soil underlying snow is frozen, which is the most case, it is very easy to detect it. But if the underlying soil is thawed or the snow is melting (as we are using observations during night, this is very accidental case), there could be certain detecting error. Unfortunately, at current stage there have not effective approaches to dealing with extremely particular issues.

Changes:

We have changed the method for trend tests by using the Mann-Kendall's tau-b test combined with Theil-Sen's slope to recalculate the temporal trends. Significant changes in frost days are found in north and central parts of North America, and North Europe. As you suggested, the snow cover may be one of the reasons that we are having low accuracy in the middle latitudes. And this has been discussed in our revised manuscript.

Comment (6):

An analysis of the influence of a spatially/temporally changing snow cover on the F/T cycles is completely missing - yet, it has a very significant influence on seasonal and permanently frozen ground, as has been demonstrated by numerous papers - including papers about the Qinghai-Tibetan Plateau.

Response:

The analysis of the influence of changing snow cover on the F/T cycles is out of the scope of our study, however its effect on the detection of the F/T cycles is discussed in Section 4 as mentioned in the above response.

Comment (7):

It is often mentioned that the remote sensing data are compared with "in-situ" measurements without explicitly detailing which measurements are meant: always in-situ air temperature measurements?

Response:

Yes, In the original manuscript, it always refers to the air temperature measured at weather stations.

Changes:

In the revised manuscript, since we have conducted the comparison with both modelled soil temperature and the in-situ soil temperature from ISMN, the in-situ measurements would refer to both air temperature and soil temperature. But more detailed/specific names have been used in the revised version.

Comment (8):

Equations (1)-(3) give the discrimination values between frozen and thawed ground. But even though the authors mention several times that physically a large transition range exist between the two states, where temperature stays at or near the freezing point, but liquid water and ice content may vary as they can co-exist, no transition range in the discrimination value is given. Else, one could incorporate this transition range into an uncertainty estimation to better assess the overall accuracy of the approach. A binary cut-off at 0_C will certainly introduce many uncertainties near the freezing point.

Response:

Yes, we agree with your statement that unfrozen water and ice co-exist and vary near the freezing point. But if we have to define the freeze/thaw state, the frozen (thawed) soil should include soil starting to freeze (thaw). In addition, the freezing point of water 0C° is used, although the freezing point of soil may vary depending on its components. This treatment is commonly used in previous studies. (Jin et al., 2009, Zhao et al., 2011, Kim et al., 2011 and Zwieback et al., 2015).

Comment (9):

My main criticism of the global application of the approach concerns the accuracy calculation using equation (4): here it is not differentiated at all between stations that show frequently or at least some freezing days and stations that have no significant freezing at all. If the number of TT is much larger than FF (or if FF = 0), the accuracy is 100% or close to it by definition and does not explain anything at all regarding the accuracy of your approach.

Without excluding the TT cases in regions where no freezing occurs, a global analysis of F/T accuracy will look much better than it really is.

Response:

We quit agree with you that a global evaluation would inflate the accuracy.

Changes:

Since permafrost and frozen ground mainly exist in North Hemisphere from 26°N to 90°N, including the Tibetan Plateau (Zhang et al., 1999), we have reduced the scope of our evaluation to the area of (26°N-90°N, 180°W-180°E) when we conduct the comparison of F/T map with air temperature, modelled soil temperature and in situ soil temperature. The comparison with the GLDAS modelled soil temperature at 0-10cm depth shows an agreement of 87.6% and 89.74% at ascending and descending time respectively. And comparison with the in situ 0-5cm soil temperature from ISMN showed an overall agreement of 86.62%.

Comment (10):

page 6, line 7: "...to exclude some outliers..." - to unspecific: how many outliers were excluded by this? are there systematic patterns in the outliers?

Response:

A robust statistical algorithm was used to eliminate the influence of outliers based on the conception of distance filter (Cook, 2000) which is also successfully applied on the previous calibration of AMSR-E and TMI (Parinussa et al., 2012). Those outliers are considered to be linked to heterogeneous land cover, coastal lines and instantaneous precipitation events. 99% of brightness temperature pairs are remained and then the linear regression model is used to calculate the difference caused mainly by the calibration procedure. These have been clarified in the revised manuscript.

(11) What is the uncertainty/accuracy of Eqs (5)-(8) or in other words, how large are the residuals?

Response:

The residuals are quite small, and detailed information can be found in the following two statistic tables.

Table 2. Comparison between AMSR-E and AMSR2 before calibration

band	*RMSD*	*Bias*	*R-square*
18.7H	4.158	-0.334	0.98232
18.7V	3.37	-0.284	0.98483
36.5H	5.512	-2.843	0.97681
36.5V	4.572	-2.642	0.98106

Table 3. Comparison between AMSR-E and AMSR2 after calibration

band	*RMSD*	*Bias*
18.7H	4.135	-0.009
18.7V	3.134	0.02
36.5H	4.729	0.016
36.5V	3.734	0.065

(12) page 6, line 29-32 and page 7 first paragraph: see comment above: the accuracy values are meaningless in my opinion if all data are evaluated, as too many parts of the data set exhibit no freezing. Compare Figures 3 and 4 with Figure 8 and you will see that large parts of your global data sets have no frost days. These should be excluded from the accuracy analysis as they will incorrectly bias your accuracy values towards high values.

Response:

Accepted. Please see response to Comment (9).

(13) page 7, line 10: that's absolutely true, it will very often be the case! That's why I am surprised that you validate the F/T signal with air temperature and not with soil temperature.

Response:

Yes, it is very difficult to find a wide distributed soil temperature database at a global scale. in this study, we selected the near-surface air temperature from weather station network as a reference to mainly confirm the continuity of the F/T dataset. This dataset has also been used for F/T detection algorithm development in previous studies (Kim et al., 2011; Zwieback et al., 2015). However, we fully agree with that our evaluation is not sufficient or convincible, so that we have included additional validation against with both GLDAS modelled and in situ measured soil temperature from ISMN. Please also see response to Comment (9).

(14) page 8, line 9-10: you assume that the descending time is approx. the time of minimum temperature. Even if that would be true, it does not tell you whether it stayed below 0_C during the full 24h. Your satellite frost days (Equation 9) are therefore something totally different than the frost days defined above.

Response:

Frost day is the day with a minimum temperature below 0 ℃. We assumed that the descending time is appropriate to the time of minimum temperature, so the observation at descending will be very close to the minimum temperature.

(15) page 8, line 16: I do not see where these numbers are "coinciding"?

Response:

Maybe the word "coinciding" is not very appropriate, but if non-frozen days is around 177 days then the frozen days is around 188 days, the frost days is around 214 days. As both have a large standard deviation from 48 to 70, we think these two numbers are

“comparable”, which is used in the revised version.

(16) sections 3.4, 3.5 and 3.6.1: parts of these sections are more a general discussion than a result. Many statements in these sections are quite trivial (e.g. that frost probability depends on latitude and season) - here, and regarding the many figures, there is large potential for shortening. Without providing a real science question or an application, the statistics provided are a bit meaningless.

Response:

Thanks for your suggestion. Considering the figures are excessive, we exclude the probability of frost from this article for shortening. And some parts are consolidated including pictures. Please also see response to John Kimball’s comment (1).

(17) Why do you not use soil temperature data or active layer data from the many studies that were published about permafrost evolution/distribution on the Qinghai-Tibetan Plateau? These data are available, and I see no reason why you don’t use them and show them to compare your data (Figs 11-17). Similarly, why do you not use the snow information which is available for this specific region?

Response:

The algorithm was validated over the Qinghai-Tibetan Plateau with 4-cm soil temperature from CEOP-CAMP project in previous study (Zhao et al., 2011). And the soil temperature of active layer along the highway is from 1995-2006 (we do not have the newest data especially for the year of 2013), which is not long enough for checking its continuity of the satellite dataset.

Changes:

In the revised manuscript, we further included soil temperature data from ISMN, and among those there are two networks over the Tibetan Plateau are used.

(18) page 9, line 34-35: "This may be associated with climatic anomalies in these summers": didn’t you check the climatic data of these summers? That can easily be verified and the corresponding sources be found.

Response:

Thank you for your suggestion. We have reviewed literatures and corresponding sources and found that the temperature in summers of 2005 and 2006 have a significant increasing trend (Wu and Zhang, 2008) which is coincided with our results. But we do not find evidence regarding to the temperature in summer of 2010.

Changes:

This is explained and cited in the revised manuscript. The sentence has been changed to “The changes of area fraction are probably associated to the climatic anomalies in these summers.”

(19) section 3.6.3, page 10, line 14-15: "57.16% and 12.03% of the area begins to freeze in July and August...": I doubt whether the 57% in July is correct, it much more probably is due

to STILL FROZEN ground conditions in July. Depending on snow conditions and altitude, the active layer in permafrost starts to thaw in late spring/early summer, so it probably is still frozen in the beginning of July at higher altitudes.

Response:

We defined that the first date of frost starts from July 1st. It is true that there are some areas is still frozen in the beginning of July because they could be frozen all year around and these areas are included. 57.16% area is based on the satellite's observation, which is only sensitive to the top layer of soil.

(20) page 10, line 34: the term "near-surface" is used throughout the text, but it is not always clear (e.g. here) whether you mean atmospheric near surface or soil near surface? That has to be clarified.

Response:

It means the near-surface soil. It has been clarified in the paper title and also in the text.

(21) page 11, line 12-13: "validated using in situ 4-cm soil temperature" –> that's the first time this important fact is mentioned !! to which figures does that apply? How many stations are used? Where are they located? These details are important to assess the real accuracy of your approach!

Response:

The algorithm was validated over the Qinghai-Tibetan Plateau with 4-cm soil temperature from CEOP-CAMP project in previous study. This is available in Zhao et al., 2011, and it would not be detailed in this paper.

(22) The conclusion is a mixture of repeating out-of-context statistical values and statements which refer to aspects which have not been discussed in the manuscript, e.g. The ice content and related sensitivity of permafrost to climate change. Discussion and Conclusion have to be better separated.

Response:

Thank you for your suggestion. The structure has been re-organized especially the discussion and conclusion have been separated.

(23) Figure 3 and Figure 4: these figures are misleading, as there are no or little frost days in many regions on this map. Trivially, the accuracy is 100% or close to 100% there. Both maps are rather maps of frost days than real accuracy maps.

Response:

Accepted. These figures have been replaced by new figures just focusing on the domain of the permafrost and frozen ground in North Hemisphere.

(24) Figure 5: what are the error bars? Explain in the caption.

Response:

Accepted. The error bars are 3 times of the standard deviation of the value in the 0.05 step of X-axis and has been clarified in the manuscript.

(25) Figure 6: which line is which year (green / black). Explain in the caption.

Response:

The black line is 1:1 line and the green line is the linear regression line. It has been clarified in the caption.

(26) Figure 8: here, the large area without frost days become apparent. The analysis should be restricted to the areas with frost days, i.e. basically the mid- and high-latitude northern hemisphere.

Response:

Accepted. The research domain has been modified to the area of 26°N-90°N.

(27) Figure 10: What is the word "water" above the color scale referring to? The location of the Qinghai-Tibetan Plateau should be indicated somewhere.

Response:

It refers to water bodies. And this figure has been excluded.

(28) Figures 11, 12, 16 and 17: the color scales cannot be read!

Response:

Accepted. These color scales have been fixed and its font size has been increased in the revised version of manuscript (figure 4 and 5).

Figure 4. The frost days and its standard deviation on Qinghai-Tibetan Plateau

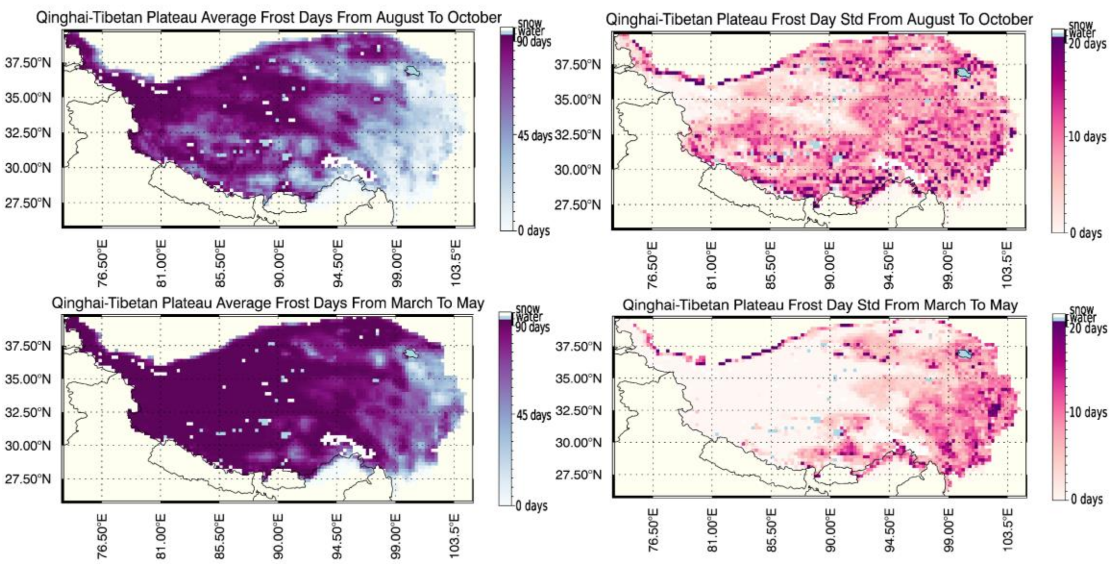
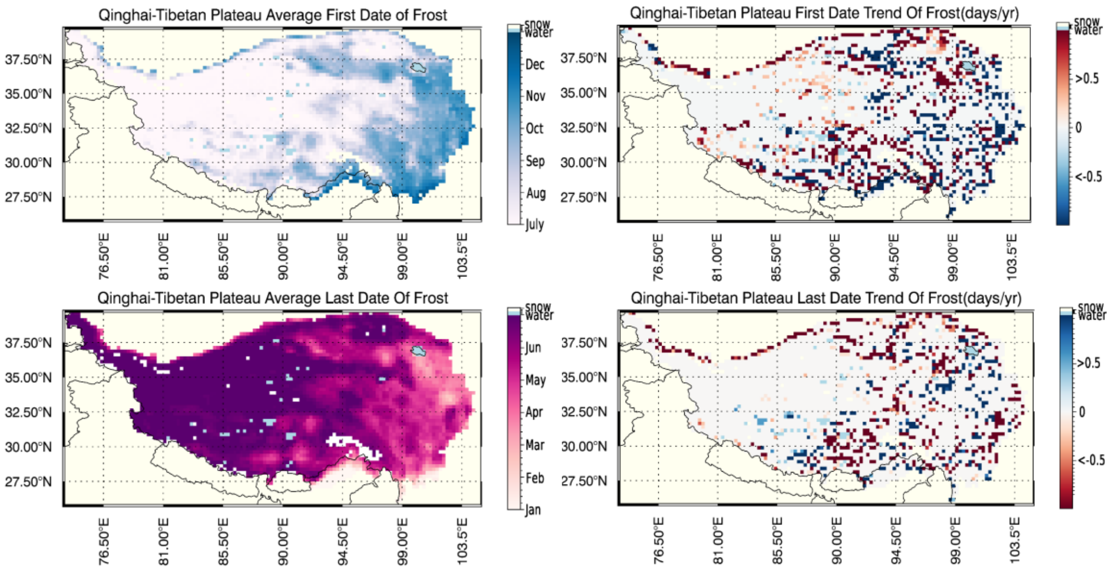


Figure 5. The frost days and its standard deviation on Qinghai-Tibetan Plateau



(29) Figure 14 and 15: the format of the x-axis has errors.

Response:

As you suggested, figures are excessive in this paper and those results has been summarized in text only.