tc-2016-115-SC2 by J. Kimball (johnk@ntsg.umt.edu)

Dear Prof. John Kimball,

Hereafter the responses to your valuable comments.

(Note: In all the replies, the number of figures and tables are not coincided with that in the revised manuscript. It is just ordered in supplement file)

General Comment:

This paper presents a new satellite passive microwave based detection of estimated near surface soil freeze/thaw (FT) status over a global land domain derived using calibrated AMSR-E and AMSR2 18 and 36 GHz brightness temperature records. The FT data records were derived separately for ascending and descending orbit brightness temperatures and validated with in-situ surface minimum and maximum daily air temperatures from global in situ weather stations. The resulting analysis includes a global accuracy assessment and quantification of global frost probability, annual frost day duration, timing of first & last frost dates, and associated trends. As a case study, regional patterns of mean frost days and trends over the Qinghai-Tibetan Plateau (QTP) was analyzed.

The paper covers a topic that is suitable to readers of The Cryosphere and should be of particular interest to those interested in climate change impacts and feedbacks relating to the changing frozen season. However, the paper suffers from a many critical weaknesses and limitations that need to be addressed before the paper can be considered to be of publication quality. These limitations are summarized below.

Response:

Thank you for the general description and your affirmation to our work. We have gone through your comments very carefully and addressed limitations that you mentioned.

Major Comment (1):

First, the paper suffers from numerous errors in the use of English grammar and sentence structure that severely detracts from the quality and reader comprehension of the material presented. The authors should enlist the help of a technical writer to revise and improve the structure and writing of the paper. The paper also needs to be restructured to clearly separate methods, results and discussion sections. For example, a comparison of the FT results against global temperature anomalies is presented in the Discussion and Conclusion section (p. 11-12), and should be moved to the Results section. The methods for deriving frost days and frost probabilities (p. 8) should be moved to the Methods section. The number of figures presented is excessive (18!) and should be cut back; Figures 6, 7, 13, 14, 15 and 18 can be eliminated as the information in these figures can easily be summarized in the text. Other Figures should be consolidated as follows: 3,4; 8,9; 11,12; and 16,17.

Response:

Thanks for your constructive advices. We have invited a native English speaker to revise the whole manuscript and its structure. And all your suggestions regarding to

restructuring have been accepted.

Changes:

The number of figures has been cut back to 13. Figure 6 and 7 in original manuscript are consolidated to Figure 6 (new version of manuscript), as well as figure 9 and 18 to figure 10, figure 11 and 12 to figure 12, figure 16 and 17 to figure 13. And Figure 13, 14, 15 have been eliminated and the information is summarized in the text (Please see section 3.5.2 in the revised manuscript).

Major Comment (2):

The authors purport to conduct effective retrievals of near surface soil FT conditions using higher frequency (18.7 and 36.5 GHz) brightness temperature records from AMSRE and AMSR2. However, validation of the FT record is conducted using in situ daily air temperature measurements at approximately 2m measurement heights from the global weather station network. Air temperature may be related to soil temperature, but the relationship may vary depending on multiple factors, including the presence and condition of snow cover, soil type and surface organic layer thickness, vegetation cover and soil moisture. The spatially integrated satellite FT signal encompasses vegetation soil, snow and other landscape elements within a coarse (25-km Res.) sensor footprint. The higher frequency (37 and 36 GHz) brightness temperature retrievals are unlikely to be directly sensitive to soil FT conditions in most areas due to the rapid extinction of microwave emissions from overlying snow, vegetation, and moist soil litter and surface layers (e.g. Du et al. 2014). Additional validation of the FT retrievals against in-situ soil temperatures is required to justify that the FT data are truly detecting soil thermal and FT related changes. Although two previous studies are cited (Chai et al. 2014, Zhao et al. 2011) as justification for the detection of soil FT, limited evidence from prior regional studies is insufficient to confirm that the AMSR FT record is detecting soil FT dynamics over a global domain.

Response:

We quite agree with you that the comparison of AMSR derived F/T with air temperature is not a straight forward approach, since there could be a time lag between air and soil temperatures. It also can be the case for modeled soil temperature (ECMWF the first layer of soil about 0-7cm, GLDAS of 0-10cm) and in situ measured soil temperature, because the microwave penetration depth might vary depending on the condition of land cover. In addition, the disagreement between satellite footprint scale estimates and point scale ground measurements make the direct validation becomes impossible. Therefore, 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 so that we have included additional validation against with both modelled and in situ measured soil temperature. Regarding to the overlying snow and vegetation effects, the algorithm we used is developed from both ground-based measurements and model simulations, which include snow covered surface, vegetation covered surface and snow-vegetation both covered surface. Therefore, it is relatively

robust over complex land surface conditions. For example, in the case with snow cover, the volume scattering at 36.5 GHz is stronger than that of 18.7 GHz so that the emission at 36.5 GHz would be lower than 18.7 GHz, which results in an increase in the Quasi-emissivity even greater than 1. Therefore, the snow-covered surface is likely to be classified as frozen, which is reasonable in most cases.

Changes:

According to your kind suggestion, we have implemented the comparison (not validation) with GLDAS modelled soil temperature and in situ measured soil temperature from ISMN. Detailed descriptions of these new external data are presented in Section 2.2. As the original domain (global) of the derived F/T map is not very convincible for one reviewer so the comparison with different in situ temperature or modelled temperature was conducted only over the North hemisphere zone of 26°N-90°N. The new results found that the satellite derived F/T map has an agreement 82.25% (Spatial agreement) and 85.43% with WMO air temperature in 2010 and 2013 respectively. The comparison of F/T map 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 (Figure 1 in supplement). As comparison with the in situ 0-5cm soil temperature from ISMN, we selected 220 stations located in the north of 26°N and the overall agreement is 86.62% (Figure 2 and Table 1 in supplement). These new results are added into the Section 3.2.

Figure 1. Comparison of F/T map with soil temperature modelled by GLDAS

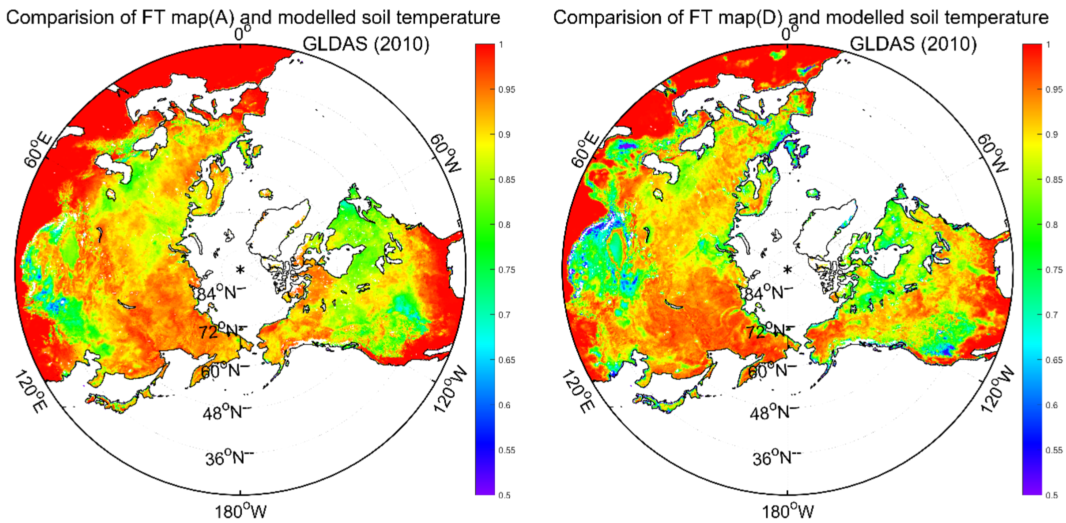


Figure 2. Comparison of F/T map with soil temperature from ISMN

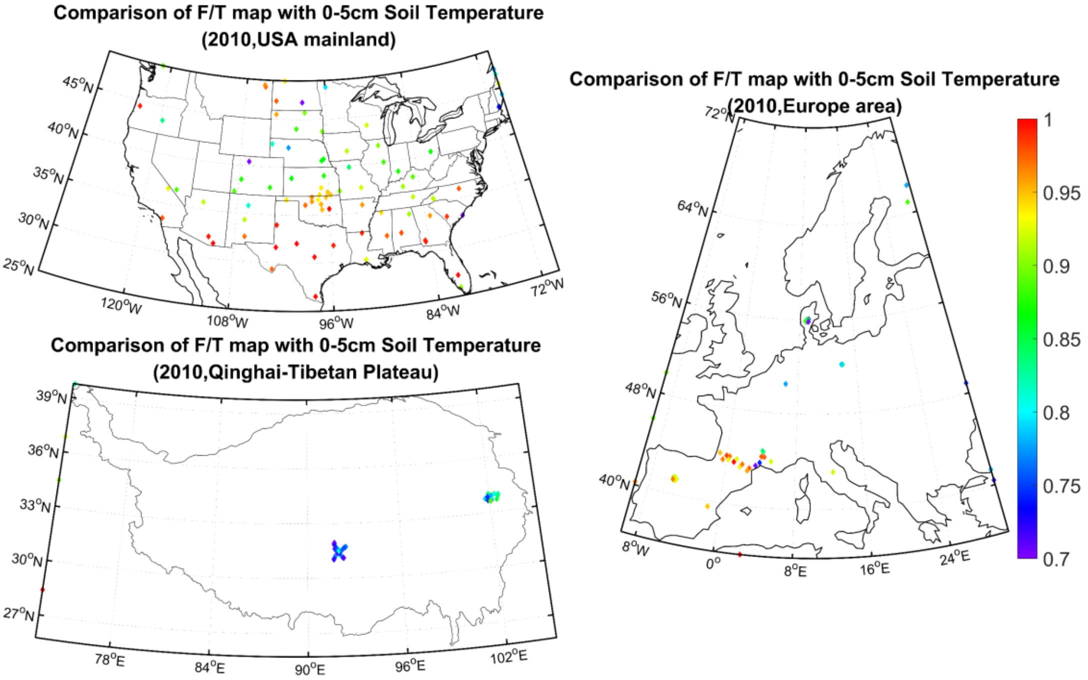


Table 1 The main soil moisture networks distribution (ISMN) and the agreement with F/T map

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		

Major Comment (3):

The authors use combined similar brightness temperature (Tb) records from AMSR-E and AMSR2 using an empirical calibration between overlapping AMSR-E Slow Rotation Mode (Level 1 S) data and AMSR2 standard data. This is potentially a very interesting approach and analysis that would be of much interest to the community; however, the approach as currently presented lacks sufficient detail and requires more complete methods development and documentation. The post-2011 AMSR-E Slow Rotation data are used to calibrate and match AMSR2 data to the prior AMSR-E operational record (2002-2010); However, this calibration assumes consistency between the AMSR-E operational record and Level 1 S data, which has not been demonstrated. The AMSR-E operational record and L1S data are not the same; the authors need to first demonstrate that the global relationship between AMSR-E and AMSR-E L1S data is consistent before using the L1S data to calibrate AMSR2 to the AMSR-E operational record. Methods Section 2.3.1 describing the AMSR inter-calibration is also lacking sufficient details on the specific versions of AMSR-E and AMSR2 data used in the study. More information is needed in Section 2.1 describing the AMSR channel frequencies available, native footprint resolutions and polarizations, and the specific channels used for the study.

The empirical Tb calibration equations (Eqns. 5-8) fail to denote whether the same equations apply to ascending or descending orbit Tb records. More information is also needed in the Results section (3.1) on whether there was evidence of regional or seasonal bias in the Tb calibration relationships, since inconsistent bias is more difficult to eliminate with single point calibration and may contribute to artifacts influencing consistency in FT classification accuracy and regional trends.

Response:

There are different approaches for the inter-comparison of AMSR-E and AMSR2. One approach is to use another microwave instruments such as the TRMM TMI and FY-3 MWRI, which can overlap the observation periods of both AMSR2 and AMSR-E. However, because of the difference in sensor configurations, corrections have to be done well with these uncertainties. Another more direct approach is the statistical inter-calibration that uses the AMSR-E slow rotation observations, which are made right for the inter-comparison with AMSR2. This approach is best suited for the two instruments having simultaneous observations. Regarding to the consistency issue between normal rotation 40-rpm and slow rotation 2-rpm data, it has been confirmed by studies from JAXA (Imaoka et al., 2016). In this study, we only focus on 18.7 and 36.5 GHz over the land, as these data are used for the F/T detection in our algorithm.

Changes:

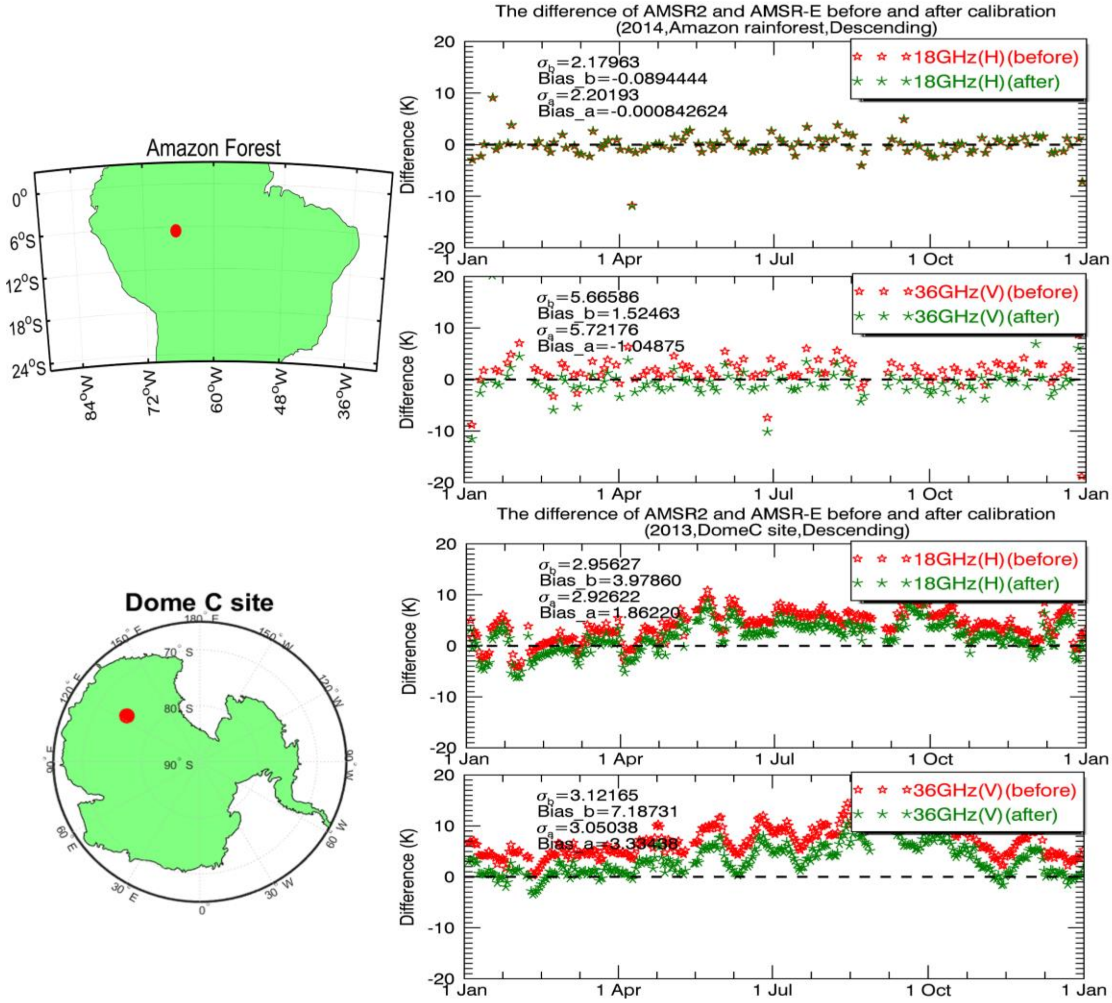
According to your comment, we have added the above descriptions to Section 2.3.1 to eliminate the reader's concern. The data version for AMSR-E is from NSIDC (Version 03), and the slow rotation data is Version release 4 obtained from GCOM-W1 Research Product Distribution Service. AMSR2 brightness temperature is extracted from Level 3 products (Version 01). These information is well described in Section 2.1 together with a Table (Table 2) of information of frequencies, footprint sizes and polarizations. In addition, we also checked the regional bias over reference targets of Antarctic ice sheet and Amazon rainforest (Figure 2). It is found that the bias of difference (ASMR2-AMSR-E) at Dome C site for 18 and 36.5GHz and at Amazon rainforest at 36.5GHz has reduced largely before and after calibration of AMSR2. While it does not show reduction at 18GHz at Amazon rainforest (actually the differences before and after calibration are very close), that is may because the observation of AMSR2 at Amazon rainforest are very close to that of ASMR-E (It is can be seen the bias are very small no matter it is calibrated or not). Additionally, only 108 data pairs are available for the same day in a year because of the gap of satellite observations. The Dome C site doesn't experience the gaps so the result it shows will reflect much of effects of the calibration models. The variances for both sites don't change much at different periods of a year. It lends much support that little seasonal bias was introduced by these models.

Table 2. The main specifications of AMSR-E and AMSR2

	AMSR-E	**AMSR2**
Satellite Platform	GCOM-W1	AQUA
Altitude	705km	699.6km
Observation Frequency (GHz)	6.9, 10.65, 18.7, 23.8, 36.5, 89	6.9/7.3, 10.65, 18.7, 23.8, 36.5, 89
Swath Width	1450km	1450km
Antenna Size	2m	1.6m
Equator Crossing Time (Local Time Zone)	1:30±15min PM Ascending 1:30±15min AM Descending	1:30±15min PM Ascending 1:30±15min AM Descending
Polarization	H & V	H & V

Note: H: Horizontal. V: Vertical

Figure 3. The Brightness Temperature changes at Dome C and Amazon



Major Comment (4):

The authors use a parametric linear least-square regression analysis to evaluate FT related trends and their significance. However, temporal autocorrelation and non-normal data distributions in time series data can lead to inflated correlations and significance using linear least squares regression analysis. A non-parametric trend analysis approach should be used for the FT trend analysis, rather than linear regression. Suitable methods for FT trend analysis include Kendall's tau with supporting tests for data normality and temporal autocorrelation of the data (e.g. see Kim et al. 2012).

Response:

Thank you for the nice suggestion, 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.

Changes:

Results in Section 3.4 and 3.5 are updated. Related figures including Figure 10 and 13 in manuscript but is showed as Figure 3 and 4 in supplement.

Figure 4. The anomalies of the number of frost days trend (days/month) and global land temperature trend (C°/month).

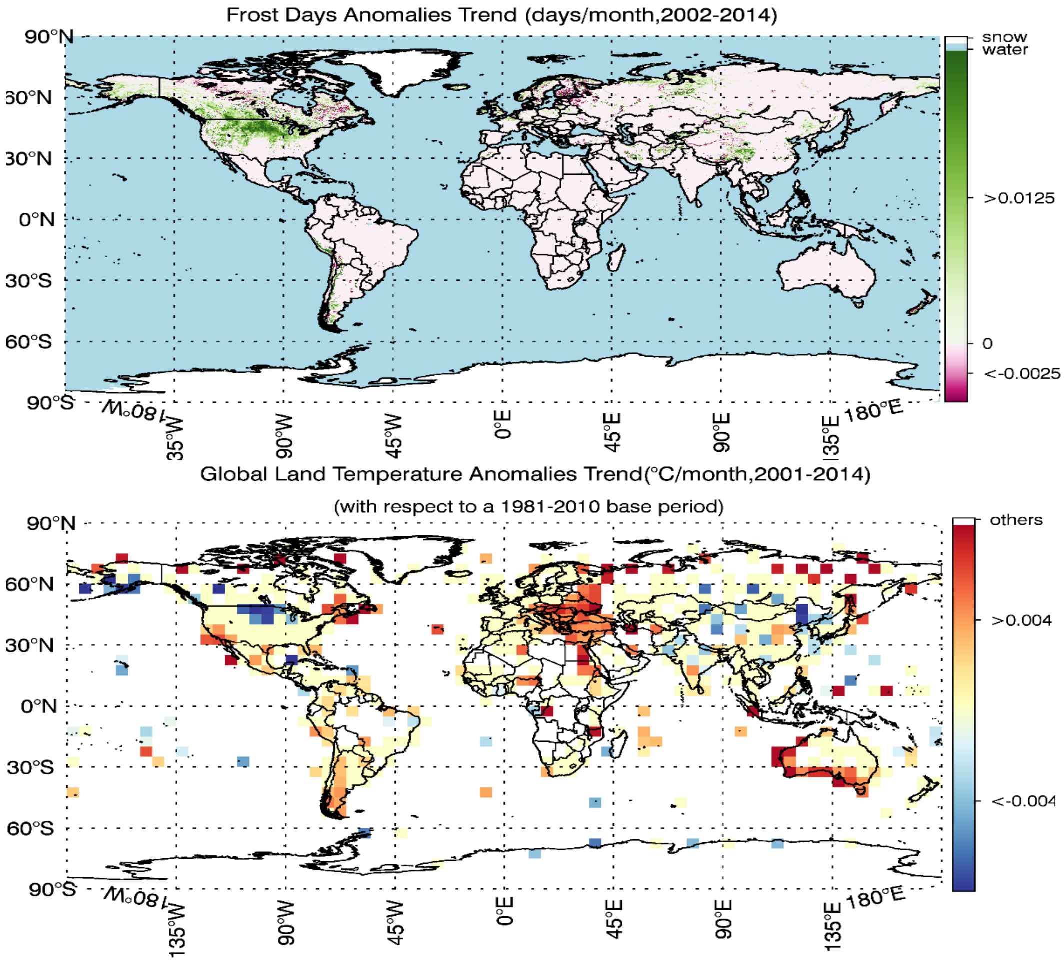
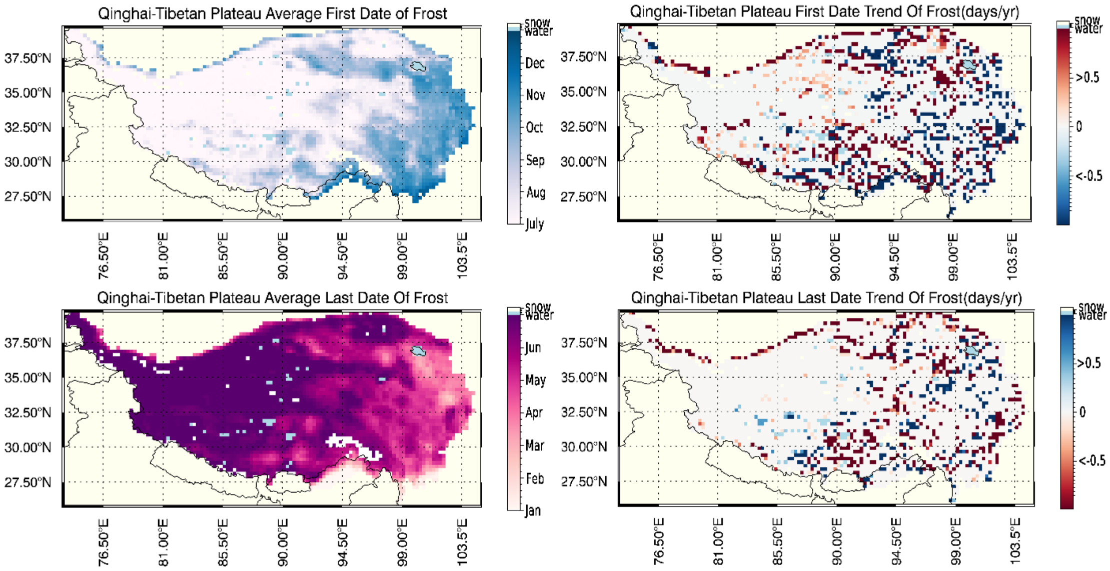


Figure 5. The trend of Qinghai-Tibetan Plateau average first and last date of frost, and their changing trends.



Other more minor comments and recommended changes are noted below.

P. 3, Ln 9: missing year in Mladenova et al.

Response:

Accepted. Thank you for your reminding, the year has been added.

P.3, Ln 10: Be careful when using the term 'first'. Inter-calibration of AMSR-E and AMSR2 has been conducted in prior studies (e.g. Du et al., 2014).

Response:

Accepted.

Changes:

'first' has been deleted. The Reference of "Du et al., 2014. Inter-calibration of satellite passive microwave land observation from AMSR-E and AMSR2 using overlapping FY3B-MWRI sensor measurements. Remote Sensing, 6, 8594-8616" has been added.

Line31, pg3: provide data source of AMSR-E and AMSR2 precipitation product (e.g., website, paper).

Response:

Accepted. The website of the data source has been added.

P. 4, Ln 6, Suggested sentence revision for better clarity: "A previous study successfully used similar in situ air temperature measurements to evaluate the daily F/T classification accuracy from SSM/I (Kim et al., 2011)."

Response:

Accepted. The sentence has been revised according to the suggestion.

P. 5, Ln 2-4: Include more information on the masking technique used in this study. Authors should include the details on outlier detection along coastal lines and large water fraction areas.

Response:

More information on the linear regression has been added (see Methods Section 2.3.1). All concurrent datasets were generated after the masking of outliers based on a standard Cooks distance filter. In addition, the data density plot provides us a possible way to filter more outlies which are considered to be heterogeneous land cover, coastal lines and instantaneous precipitation events.

P. 5, Ln 9: what is Tb37V? Is it 37GHz and vertical polarization?

Response:

Sorry to confuse you. And yes, Tb37V is brightness temperature value of 36.5 GHz at vertical polarization. It has been clarified in the manuscript.

P. 5, Ln 19-21: More information is needed here to clarify that the discriminant function was used with SSMIS and AMSR-E rather than MODIS data.

Response:

The discriminant function algorithm is developed based on the AMSR-E sensor. Its accuracy has been previously assessed by comparing with MODIS land surface temperature product and *in situ* 4-cm soil temperature. This has been clarified in the revised manuscript.

P. 5, Ln 30-31: Be more specific here to distinguish whether the two FT classifications indicated are from in situ air temperature and AMSR data, or AMSR ascending and descending orbit data.

Response:

Accepted.

It has been specific that the two classifications are for AMSR ascending and descending orbit data.

P. 6, Ln 32: The authors state that the AMSR FT classification results "are better or comparable with previous studies (Kim et al. 2011)". However, the Kim et al. study documented a mean annual FT classification accuracy over a much smaller global domain and different period than the current study. The Kim et al. study was limited to vegetated land areas where seasonal frozen temperatures are common and represent a significant environmental constraint to ecosystem productivity. The current study encompasses nearly all global land areas, including the warmer tropics where frozen temperatures never occur, which leads to a somewhat inflated FT accuracy metric (e.g. Fig. 4). It would be more appropriate to compare FT accuracy over a domain more consistent with prior studies (e.g. land areas above 45N as reported by Kim et al., 2014). Otherwise, a more appropriate statement here would be that the AMSR FT classification results are "similar" or "comparable" with previous studies.

Response:

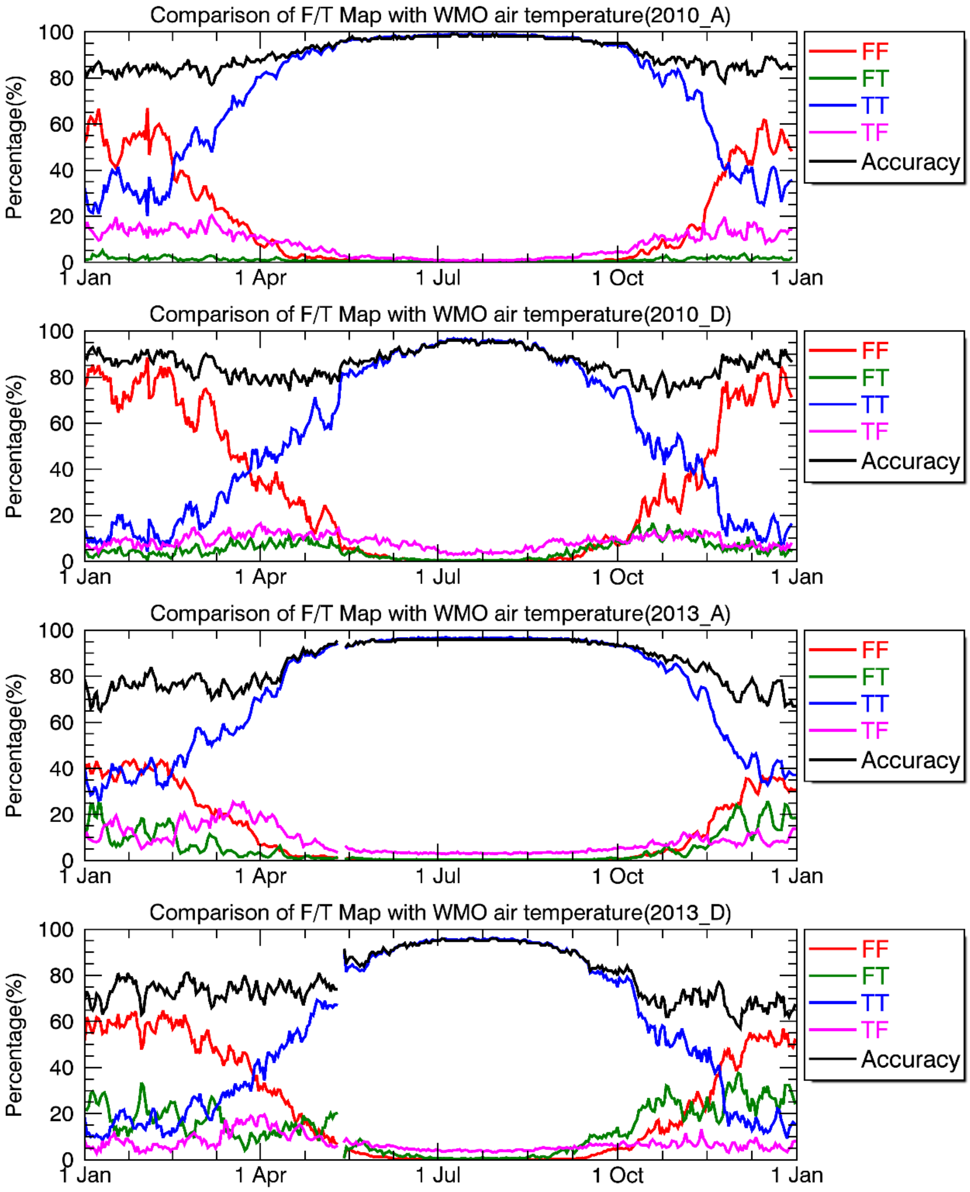
Accepted. The domain has been reduced to the area of (26°N-90°N), which is consistent with the statistics of permafrost (Zhang et al., 1999). The original sentence has been revised by a more appropriate statement "This is similar or comparable with previous studies (Kim et al., 2011)" as suggested.

P. 7, Ln 9-11: The result shows a different pattern in descending orbits. There is a lower percentage in TF, 2013 at descending overpass (fig2).

Response:

As there are comments suggest that a global land domain would lead to inflated F/T accuracy metric, we redefined our research domain (see above responses). The new results also show that the general temporal pattern of different percentages of FF, FT, TT and TF in 2010 and 2013 are similar (See Figure 5). However, it is true TF percentage in 2013 at descending orbits shows a different pattern which is lower than FT during the winter. This could be attributed to many factors, one of explanation is that the air temperature we used as reference is generally lower than soil temperature during night (descending orbits), which would lead to an increase of FT percentage and decrease in TF percentage. And this is more significant in the year of 2013. Additionally, the in-situ air temperature we used for comparison also varies daily and has an inter-annual variability. On a given day, the global observations we selected are depending on whether the observation is available or not over each station, instead of a fixed number of stations for the whole year. Thus, for different days even the same year, the number and locations of in situ measurements could vary. We have explained these in our revised manuscript.

Figure 6. Comparison of F/T map with WMO air temperature (2010 and 2013)



P. 7, Ln 27: Include the number of stations.

Response:

Accepted. The number of stations that both used in the two different years are not fixed and it all depends on whether data records are available or not on a given day for each station. Although the exact number of stations varies day by day, but it is about 4000 which has also been included in the text.

P. 8, Ln 12: I think you mean “given day” rather than “special day” here and elsewhere.

Response:

Thank you. We are intended to use the “given day” instead of “special day”. However, according to your comment (1), the manuscript is reorganized and compressed that the frost probability map is no longer included here.

P.8, Ln 17: The value of 177.6+/-47.6 days was used in Kim et al., 2014, which may be more appropriate here.

Response:
Accepted. It has been corrected.

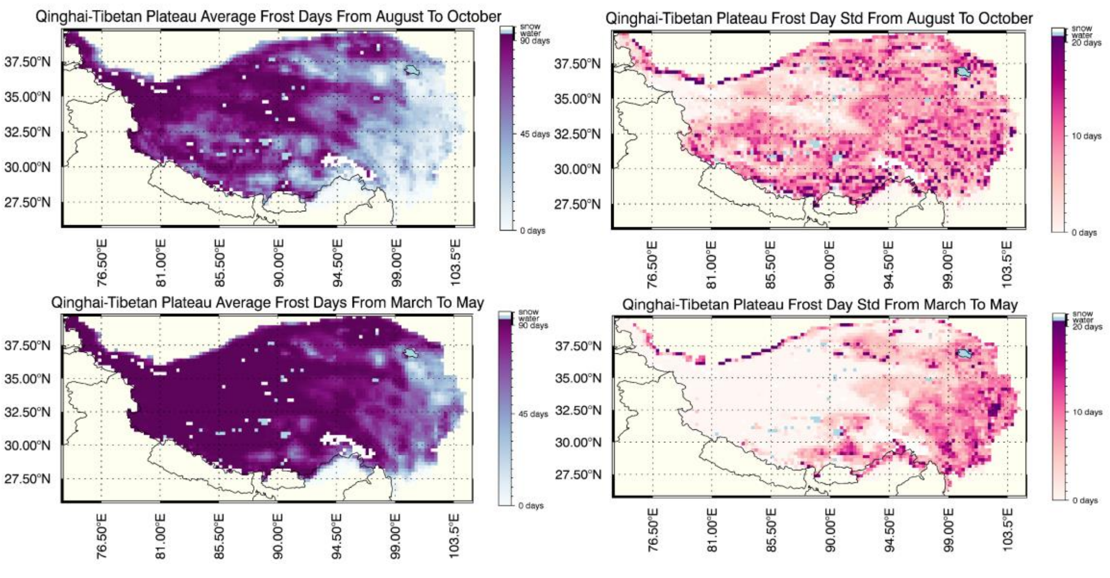
P. 9, Ln 20: Clarify whether this refers to the spatial or temporal standard deviation.

Response:
Thank you for your kind suggestion. The frost probability part in our revised manuscript has been excluded.

P. 9, Ln 25-26: Are the two standard deviations (6.04 and 6.96 days) significantly different?

Response:
They are not significantly different but still different. It is more clear in Figure 6 that that the areas with larger frost days have a smaller spatial deviation, which means the frost days in these areas are much stable.

Figure 7. The frost day and its standard deviations on QTP



P. 9, Ln 34: Summer is defined as July to September here, but fall (Ln 22) is not defined, but should be for consistency.

Response:
Accepted. We have replaced the season with the specific month period in the revised manuscript for consistency.

P. 9, Ln 35: What do climatic anomalies in summer refer to? Are they warmer summers?

Clarify.

Response:
Thank you. As the minimum area fraction of frost and the frost days during these periods (from July to September) of 2005, 2006 and 2010 are much lower than the same periods of other years. It may indicate that these periods are warmer, which is confirmed by other studies (Wu and Zhang, 2008). It has been clarified in the revised manuscript.

P. 11, Ln 13: Clarify which overpass (AM or PM) is used for the accuracy assessment (Zhao et al., 2011)?

Response:
The accuracy assessment conducted in previous study used AMSR-E ascending data at 13:30 local time and it has been clarified.

P. 12, Ln 8: Permafrost types in QTP are not discussed elsewhere in the paper. Including QTP permafrost classification (e.g., map) would strengthen the discussion regarding statements linking FT results to permafrost and active layer conditions.

Response:
Accepted.

Changes:
We have added the discussion associated with permafrost types in QTP in the revised manuscript (please see the discussion and conclusion section). The map of permafrost types is provided by Cold and Arid Regions Science Data Center (Figure 7).

Figure 8. The types of permafrost on Qinghai-Tibetan Plateau