Interactive comment on "Development and analysis of a continuous record of global near-surface soil freeze/thaw patterns from AMSR-E and AMSR2 data" *by* Tongxi Hu et al.

J. Kimball

johnk@ntsg.umt.edu

Review of T. Hu, T. Zhao, J. Shi, et al. paper titled, "Development and analysis of a continuous record of global near-surface soil freeze/thaw patterns from AMSR-E and AMSR2 (tc-2016-115).

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.

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.

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 AMSR-E 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.

The authors use combine 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.

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).

Other more minor comments and recommended changes are noted below.

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

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).

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

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)."

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 coastlines and large water fraction areas.

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

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.

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.

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.

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).

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

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

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.

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

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

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

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

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

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.

References 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 Kim et al. 2014. Attribution of divergent northern vegetation growth responses to lengthening non-frozen seasons using satellite optical-NIR and microwave remote sensing. Int. J. Remote Sens. 35 3700–21