Review of Lv et al.
Simon Zwieback (Referee)

Lv et al. present a new remote sensing algorithm for identifying the land surface freeze/thaw state using SMAP passive microwave observations. The central premise of the algorithm is that landscapes that remain frozen on daily to synoptic time scales are characterized by small diurnal differences in brightness temperature. The authors compare the results obtained with their algorithm with the SMAP product and various reanalysis-based freeze/thaw-related temperature indices over the Northern Hemisphere. They further derive common F/T metrics such as the length of the frozen period, but they do not analyze them in depth.

The manuscript's content is relevant to The Cryosphere, as it covers a topic that is of interest to the journal's audience and also to applied remote sensing scientists. However, it suffers from several major weaknesses and inconsistencies that will require extensive revision or rewriting before publication. The main shortcomings are:

1) opaque writing that curtails comprehension (paper structure, poorly structured paragraphs, figures)
2) inconsistencies between the author's claims that their algorithm is globally applicable and the assumptions underlying the algorithm
3) multiple claims that are not backed up by evidence or references
4) limited scrutiny of the algorithm and its output

1) Lack of clarity
The manuscript is difficult to read. In addition to many poorly worded phrases that detract from the content, the paper structure, the paragraph structure and the figures are challenging to follow.
The paper structure is unusual in that the authors do not include a proper discussion (see below). Furthermore, the introduction does a poor job of conveying the main ideas and findings. For instance, the variance-based filtering that is central to the algorithm is not mentioned. The reader is caught by surprise in the methods section. More broadly, I suggest the authors clarify what they mean by the estimand, i.e., the freeze/thaw state. Is it defined instantaneously (with the variance-based filtering just a convenient means to stabilize the estimation), or is it aggregated on daily or synoptic time scales?
The paragraphs often appear to be haphazardly put together, thus greatly limiting the readability of the manuscript. The introduction serves as a good example. The paragraph starting at line 47 opens by highlighting the limitations of temperature-based indicators for freeze/thaw state estimation. The second sentence states that "[i]n contrast, more direct state information results from the very different microwave dielectric constant for frozen and unfrozen soil. However, the reader has to guess that this sentence and the paragraph it is contained in are about microwave remote sensing of the freeze/thaw state, as the expressions "remote sensing" and "freeze/thaw" are not mentioned once. The remainder of the paragraph talks about emissivities without referring to the frequency and polarization. At some point, the reader stumbles upon L-band observations from SMAP and SMOS, which, however, are of central importance to the manuscript and to the introduction. I suggest the authors identify one theme for each paragraph and structure the paragraph such that the reader can easily follow.

Another example of opaque writing is furnished by lines 177-190. The authors first propose their own definition of the beginning/end of "the annual freezing", but the subsequent algorithm is seemingly at odds with the definition. For instance, a brief cold spell in summer would meet the definition but would in most cases be screened by the variance criterion. The authors further claim that their variance screening using a window length beta of 7 (implicit unit: days) "optimally" filters out brief events. However, it is not at all clear how optimality is defined and what the evidence is that optimality is achieved. The figures are exceedingly difficult to interpret. For example, figure 3 is a cornucopia of lines and markers, with poor contrast (the yellow line is almost invisible) and most items being obscured by others that are plotted on top. The choice of colours (rainbow scale) and line weights (the grid obscures a good part of Fig. 11) is questionable in almost all maps. The captions contain insufficient information to interpret the figures. For instance, Fig. 11 does not explain how the fraction of agreement (negative in the figure) was computed. Is it a difference? The caption of figure 2 shows a histogram of the beta-windowed variance, but it is not explained what input data were used and what value of beta was used.

2) Globally applicable algorithm?
The authors claim in the title that their algorithm is globally applicable, but the limited applicability of the underlying assumptions casts doubt on this claim. The authors do little to dispel these concerns, as they do not include a separate discussion section where associated limitations ought to be scrutinized. I also note that despite the word global in the title, no results for the southern hemisphere are provided. However, my two biggest concerns in this respect are the assumption of 6 am / 6 pm overpasses and the assumption of elevated variability of the dielectric characteristics of thawed landscapes on synoptic scales. Neither of these two assumptions are very accurate on a global scale. The assumption of 6 am / 6 pm overpasses is difficult to defend at high latitudes, where the temporal sampling deviates substantially from that at the equator. The authors neglect this issue completely, although negative repercussions on their algorithm's performance are not too difficult to imagine. The assumption of elevated variability of the dielectric characteristics of thawed landscapes on synoptic scales is not subjected to any scrutiny. The authors acknowledge that the Rossby wave time scale that serves as foundation for the beta parameter is relevant to mid-latitudes, but they do not discuss their variance-based filtering within a time window of length beta affects the results elsewhere. Among the regions of particular concern, I list cold arid regions (mentioned by the authors as presenting challenges to microwave F/T algorithms in the introduction), bedrock-dominated areas, and regions with extended periods of stable anticyclone.
3) Unsubstantiated claims
The authors make numerous claims that are not backed up by evidence or references. An excellent example is furnished by the paragraph starting on line 153, whose intent is to provide a rationale for the new algorithm. There, the authors make numerous such claims. For instance, they state that brightness temperature changes during freeze/thaw transitions are at least as large as those associated with precipitation "because of the huge epsilon difference between frozen and unfrozen soil". They do not provide a reference or evidence for this claim, nor do they state when this may not be the case (e.g., certain arid landscapes). A further issue is that the language is inappropriate and vague ("huge"). There are numerous similar claims in this paragraph alone, and not a single piece of evidence or reference is provided.

4) Very limited scrutiny
The authors do not subject their algorithm and its underlying assumptions to the level of scrutiny that a reader of The Cryosphere may expect. There is no discussion section that assesses failure cases or that establishes a link between potentially inappropriate assumptions and questionable results. Furthermore, general issues with the "validation" strategy employed here (e.g., scale and commensurability with reanalysis-derived temperature metrics) should be incorporated.

Minor comments
l 37: suggest replacing solar with shortwave and terrestrial with longwave
l 40: Potentially inappropriate reference (Schuur et al.): How does the surface freeze/thaw state relate to permafrost carbon
l 94: "replaying": odd choice of word
l 171: That \Delta TB_i will be smaller than \Delta T_i does not follow from the provided inequalities because (7) is a sum. A mathematically sound argument is needed to substantiate the claim.