Reviewer #1

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

Reply: Thank you very much for your meticulous comments and feedbacks. We have revised the paper and explained your comments in the following.

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.

Reply: Sorry for that. I improved the manuscript point by point.

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.

Reply: Thanks a lot for this comments. We had made corresponding changes to make this point clear:

“The new FT algorithm has a similar basis to the DAV approach(Sharifnezhad et al., 2021) applied to higher frequency passive microwave measurements for snow and ice sheet applications(Kopczynski et al., 2008; Tedesco, 2007). It has been proved the Durinal Amplitude Variation (DAV) of passive microwave signals are sensitive to the FT state changes(Sharifnezhad et al., 2021), which are dynamic and complex and vary continuously in space and time.” is shifted (revised) from Section 2.3 to Line 66-67.

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?

Reply: sorry for the unclear statement here. From Eq. 1 to Eq. 8, we took the daily scale as the other DAV approaches (added in Line 183). By adding Eq. 9 and Eq. 10, we add synoptic time scales background to daily values (added in Line 203). Some clarifications were added to make

this point clear.

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.

Reply: Thanks a lot for your comments. We'd explicitly state in our revisions in Line 47. "However, FT state estimations from in-situ temperature observations are limited in scale, and it is not straightforward to deduce the state from soil, skin, or near-surface air temperature." is shifted to the previous paragraph. From MEaSUREs to SMAP' FT products, the paragraph is about remote sensing of FT state with the passive microwave remote sensing.

The remainder of the paragraph talks about emissivities without referring to the frequency and polarization.

Reply: Yes. And "For the L-band." is added in Line 51.

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.

Reply: We revise the Introduction part, especially to ensure the first sentence indicates each paragraph's core idea. We hope it is fine for the potential readers now.

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.

Reply: Thank you for this comments. "a brief cold spell in summer would meet the definition but would in most cases be screened by the variance criterion." is exactly how the new algorithm works.

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.

Reply: Thank you for your suggestion. We agree that "optimally" is not appropriate here because Figure 9&10 already shows that it depends on the location and scale. "optimally" is deleted from the text.

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.

Reply: As suggested by another reviewer, Figure 3a removes soil temperature time-series.

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?

Reply: thank you. The caption is revised as "Figure 11: The spatial pattern of the fraction of

agreement difference compare to Figure 9b for β=3 (a), β=11 (b), γ=5 (c), and γ=11 (d)."

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.

Reply: Thanks a lot. The caption is revised as " $\mathrm{var}(\Delta TB)_{\beta}$ with β=7 by SMAP TB data contained in SMAP L3 Radiometer Global Daily 36 km EASE-Grid Freeze/Thaw State, data over the northern hemisphere where 95% of samples are within 8K."

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.

Reply: we remove the word "global" from the title. This is the only place we refer the new algorithm's adapbitlity. However, we have to note that all sites in SMAP handbook (Algorithm Theoretical Basis Document (ATBD) SMAP Level 3 Radiometer Freeze/Thaw Data Products (L3_FT_P and L3_FT_P_E)) for FT Cal/Val are taken from the northern hemisphere, and SMAP provide global FT products.

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.

Reply: it is to note that the 6 am / 6 pm overpassing time are not assumptions. In SMAP hand book (Algorithm Theoretical Basis Document (ATBD) SMAP Level 3 Radiometer Freeze/Thaw Data Products (L3_FT_P and L3_FT_P_E)) on page 26 and other places, it is said "The SMAP L3_FT_P product distinguishes 4 levels of freeze/thaw conditions determined from the ascending (6AM) and descending (6PM) orbit retrievals, including frozen (from both AM and PM overpass times), non-frozen (AM and PM), transitional (AM frozen; PM non-frozen) and inverse-transitional (AM non-frozen; PM frozen) states". In all cases, the daily L3_FT products incorporates (AM and PM satellite overpasses) data for the current day, as well as past days information (to a maximum of 3 days, necessary only near the southern margin of the FT domain) to ensure complete coverage of the FT domain in each day's product.

The assumption of elevated variability of the dielectric characteristics of thawed landscapes on synoptic scales is not subjected to any scrutiny.

Reply: NASA MEaSUREs already can distinguish FT heterogeneity in accordance with mesoscale climate and landscape, please refer to http://www.ntsg.umt.edu/freeze-thaw/

Kim, Y., Kimball, J. S., Glassy, J., and Du, J. Y.: An extended global Earth system data record on daily landscape freeze-thaw status determined from satellite passive microwave remote sensing, Earth System Science Data, 9, 133-147, 10.5194/essd-9-133-2017, 2017.

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.

Reply: Thanks for Reviewer's comments. We believe this is not the case as perceived by the reviewer. Figure 10a discusses how beta affects the domain(lat>20°N). If elsewhere means a particular place, please refer to Figure 11. Among the regions of particular concern, we 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).

Reply: Thank you. Please refer to Sharifnezhad et al., 2021 with a DAV signals analysis. And we'd added this citation in the revision.

Sharifnezhad, Z., Norouzi, H., Prakash, S., Blake, R., and Khanbilvardi, R.: Diurnal Cycle of Passive Microwave Brightness Temperatures over Land at a Global Scale, Remote Sensing, 13, 817, 2021.

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.

Reply: Thanks a lot. We try to amend throughout the manuscript and change the words like "huge" to an appropriate phrase.

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.

Reply: Thanks a lot for this comment. We try to present such scrutiny via formulas and numbers in Section 2.3. Besides, we're trying to repeat the work with SMAP's core site data for SMAP's FT products. The accuracy in SMAP's FT products is expected to be 80%, and some studies show that it is >70%. The new algorism is not validated because of the reasons explained in Discussion, i.e., lack of ground truth. What we did is to prove that the new algorithm does relate to the FT state transition in theory, and it has comparable accuracy with SMAP's. We now also include some of limitations of our theory in the discussion in Line 369-376.

Furthermore, general issues with the "validation" strategy employed here (e.g., scale and commensurability with reanalysis-derived temperature metrics) should be incorporated.

Reply: Thank you for your suggestion. We add, "The SMAP FT team uses WMO's air temperature, and WMO's air temperature is still sparse for the agreement assessment. For instance, how to deal with the scale mismatch between the weather station and SMAP's

footprint? How to account for sub-grid open water fraction, terrain heterogeneity, tree cover, precipitation, and snowmelt and on with the weather station data? These problems can be avoided for ERA5-land air temperature in the evaluation, and we are aware that the ERA5-land air temperature is not appropriate for validation which needs FT ground truth for sure." In Line 325-328.

Minor comments

l 37: suggest replacing solar with shortwave and terrestrial with longwave

Reply: These words are updated.

l 40: Potentially inappropriate reference (Schuur et al.): How does the surface freeze/thaw state relate to permafrost carbon

Reply: as mentioned in P174 as "A number of ecosystem and Earth system models have incorporated the first approximation of global permafrost carbon dynamics. Recent key improvements include the physical representation of permafrost soil thermodynamics and the role of environmental controls, in particular the soil freeze/thaw state, on a decomposition of organic carbon". We replace it with

44. Koven, C. D., Riley, W. J. & Stern, A. Analysis of permafrost thermal dynamics and response to climate change in the CMIP5 Earth system models. J. Clim. 26, 1877–1900 (2013).

45. Koven, C. D. et al. Permafrost carbon-climate feedbacks accelerate global warming. Proc. Natl Acad. Sci. USA 108, 14769–14774 (2011).

l 94: "replaying": odd choice of word

Reply: We refer to https://www.eea.europa.eu/data-and-maps/data/external/era-interim-1 for the origins of "replaying".

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.

Reply: (7) is not just a sum but a sum with weighting function. The sum of the weightings, i.e., $\left(1-e^{-\tau_1}\right)<1$, $\left(1-e^{-\tau_i}\right)\prod_{j=1}^{i-1}e^{-\tau_j}<1$ and $\prod_{j=1}^{n-1}e^{-\tau_j}<1$, is equal to 1.