Reply to comments:

However, there are still several issues that I have identified with the revised manuscript that I think must be addressed before publication is possible. These issues mostly relate to the estimation of annual deformation rates, comparison of said rates to the LiDAR-derived ΔDTMs, and model fitting of these annual deformation rates. I have summarized these points below:

-My primary remaining concerns regard the estimation of the annual deformation rate (Section 4.2.2) and subsequent model regression to estimate the recovery (most notably Figure 4). I will use the YKD case as an illustrative example because I think the potential complication is likely most pronounced here, although this comment applies to the method more broadly and hence all the study regions.

1. While the confidence interval of the fitted model provides an error envelope for the recovery time (the intersection of the model regression with the x-axis, plus or minus the 2.5th and 97.5th bounds), there is the further issue of how estimate uncertainty in the annual deformation rate propagates into the model regression. For example, in Figure 4 no error bars are included on any of the individual annual deformation rate estimates (nor thaw season deformation rate estimates for that matter); I think it is rather important to include these. Since annual deformation rates were estimated via linear regression in Section 4.2.2, there should be an associated estimate such as the RMSE or standard error that can be used.

- Reply: We thank the reviewer for this suggestion. We have added a new Section 4.5 describing the uncertainty estimates now included in the figures. We revised Figure 4 and added a new Figure A8 incorporating the error bars described in Section 4.5.
- New Section 4.5 Model and InSAR uncertainties:
    - Uncertainties of depicted data points for curve fitting are provided as error bars indicating the standard deviation across multiple results within each time-since-burn year. Additional uncertainty indicators are provided as error bars for the InSAR-derived estimates. For the thaw season deformation rate (-alphaDDT), the standard error of the linear regression between DDT and surface displacement was derived. For the annual deformation rates the standard deviations of the linear model with respect to the time-series input data are provided. For the Yukon-Kuskokwim region, where annual deformation rates were estimated using the DDT-based approach by Liu et al. (2012), the standard deviations of displacement residuals were calculated. In order to illustrate the uncertainty of the fitted functions, the 95% uncertainty envelopes are provided, similarly to Cao and Furuya (2025), estimated as the predicted value ± 1.96 times the residual standard error ($RSE = \sqrt{(RSS / (n - p))}$, where RSS is the residual sum of squares, n is the number of observations, and p is the number of model parameters). For the regionally merged result an additional 95\% confidence interval is estimated based on a parametric bootstrap (2000 samples), defined as the 2.5th and 97.5th percentiles of the resulting predictions.
- Furthermore, we added to Section 6.1:
    - The error bars representing the standard deviation across all fire scars within each time-since-burn year (Figure 4) illustrate the variability in deformation rates among individual fire scars within the same region. The InSAR-specific error bars (Figure A8) represent the uncertainty of the individual InSAR-derived deformation estimates arising from the respective estimation procedures and illustrate the confidence associated with the estimated deformation rates.

- Addition at new line 564:
    - However, the estimate reported by Michaelides et al. (2019) falls within the RSE-based uncertainty band of our fitted function for this region and is also consistent with the uncertainty of the individual InSAR-derived deformation rate estimates. This suggests that, although our fitted function indicates a faster recovery, the associated uncertainties overlap with previous observations (Figure A8b)

2. Looking at the list of interferometric pairs used in the PALSAR-2 annual deformation rate analysis, the YKD interferometric network has two sub-seasonal interferometric pairs of 42 days each, while all other interferograms used for the annual deformation rate analysis are truly close to interannual. I imagine that the authors included these two interferograms to ensure that the network remains full rank during SBAS inversion. However, there inclusion also complicates the interpretation of this time series as solely reflecting annual deformation rate, as there is implicitly a seasonal component in the solution. Furthermore, their inclusion in the interferometric network might magnify any bias present in these interferograms into the final annual deformation rate estimate, since the sensitivity to error during SBAS inversion is scaled by the inverse of the temporal baseline. For example, a 1 cm bias in a 42-day interferogram would yield a bias of ~8.7 cm/yr, whereas a 1 cm bias in a 360 day interferogram would yield only a ~1 cm/yr bias. I suspect that this in part explains the much larger dynamic range and higher spatial variability of the annual deformation rate estimate for YKD compared to the other study regions (Figure 10). The observed dynamic range of annual deformation rate is much larger than the other sites, with areas of the unburned tundra exhibiting a variability in annual deformation rate of -3 cm/yr vs +3 cm/yr over spatial scales considerably smaller than other study sites reported here (for example immediately southeast and northeast of the 2015 Kuka Creek fire, or north of the 2002 Aropuk fire). These large variations over small spatial scales to me suggest an uncompensated bias term such as a varying soil moisture efect or residual phase ambiguity due to the highly discontinuous spatial phase field -- perhaps propagated through from the subseasonal interferograms – and not merely a geodetic deformation rate. This is important as this variability will impact the statistics both within the burned and unburned reference areas, consequently impacting the 'burned-unburned' rate and the model fits and associated recovery times that are central to the manuscript. A few suggestions on how the authors might be able to address this:

   a. Are there alternative interferograms that can be incorporated into the network such that the sub-seasonal interferograms can be removed?

   b. If the answer to the above is no, can the annual deformation rate be estimated using an alternative approach, such as the stacking method (as in Chen et al. 2020: h"ps://doi.org/10.1016/j.rse.2020.112007), or by performing a linear regression to the individual interferometric pair deformation rates (as in Liu et al. 2010: https://doi.org/10.1029/2009JF001547) rather than the output time series?

- Reply: Thank you for pointing this out. We now removed the sub-seasonal interferograms. As no alternate interferograms were feasible, we used the approach described in Liu et al. (2012) in order to derive annual deformation rates for the Yukon-Kuskokwim Delta and updated the corresponding figures, tables and values accordingly. We further included following addition to Section 4.2.2:

- “For the Yukon-Kuskokwim Delta, a fully connected interferometric network could only be established by including short seasonal interferograms, as end-of-season interferograms often had to be discarded due to low coherence. To avoid the inclusion of seasonal interferograms, which would introduce a seasonal displacement component and potentially amplify any bias in the estimated long-term deformation rate, a different approach was adopted for this region. Instead, August acquisitions were incorporated, and the long-term deformation component was estimated following the approach described in Liu et al. (2012). Using similarly pre-processed and spatially filtered interferograms, the displacement *D* between two acquisition dates t_1 and t_2 was modeled as:

$$D = D_{\text{long-term}} + D_{\text{seasonal}} = R(t_2 - t_1) + E\left(\sqrt{DDT_2} - \sqrt{DDT_1}\right)$$

where *R* is the annual deformation rate, *E* is the coefficient of the thaw season deformation, and *DDT1* and *DDT2* the DDTs at t_i.”

- We have further made the following necessary changes to the manuscript:
- Section 3.4.1:
- New line 179:
  - Air temperature data used to compute degree day of thaw (DDT) dependent thaw season deformation rates (see Section 4.2.1) “and to perform DDT-based modeling of annual deformation rates for the Yukon-Kuskokwin Delta (see Section 4.2.2)” were obtained from the ERA5 reanalysis dataset.
- Section 5.1.2, new line 365:
- Old:
  - The Yukon-Kuskokwim Delta shows the largest relative heave for older scars compared to their surroundings.
- New:
  - The Yukon-Kuskokwim Delta shows on average slightly larger relative heave values for older scars compared to their surroundings.
- Added to new line 375:
  - Following the initial increased annual subsidence rate values, the Inuvik region takes the longest to align with the surrounding areas, requiring about 17 years (Figures 5b, c), “which is only slightly longer than for the Yukon-Kuskokwim Delta (about 15 years).”
- Changed new line 382:
- Old:
  - When considering ground temperatures at a depth of 2 m, fires in the region with positive ground temperatures appear to show somewhat greater deviations from the surrounding reference areas, especially for older fire scars (Figure 7b), whereas colder regions show less pronounced deviations from their surroundings.
- New:
  - When considering ground temperatures at a depth of 2 m, no clear dependency of annual deformation rates on ground temperature could be discerned (Figure 7b); however, the warmest region, the Yukon-Kuskokwim Delta, shows on average slightly larger deviations from the surrounding reference areas for older fire scars than colder regions (Figure 5b, c).

3. Figure 4: the 95% confidence intervals are now significantly tighter than in the previous version of the manuscript. The authors state in their response that “Upon revisiting the analysis, we found that the interval shown in the Figure 4 corresponded to a prediction interval, which is inherently wider than the confidence interval of the fitted model.” However, the previous manuscript version reports these intervals as confidence intervals and not prediction intervals. Some clarification is needed on this point. To avoid ambiguity and make these results as directly comparable to the prior referenced literature, I would recommend that the 95% confidence intervals that the authors report be estimated for each fitted model by reporting the predicted value +- 1.96 times the standard error, which is the method used in Cao and Furuya 2025 (note that Cao uses a parametric bootstrap to estimate the confidence interval of the integral of the annual deformation rate fitted function, but not for the confidence interval of the annual deformation rate fitted function itself).

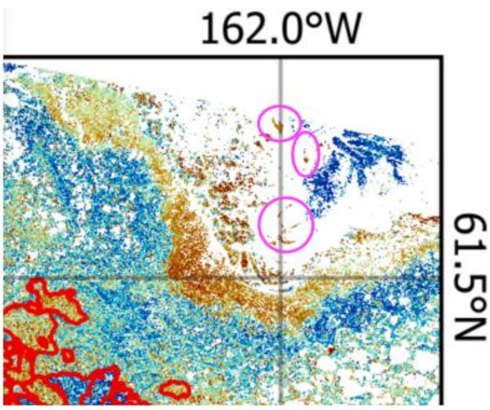
- Reply: The previous version of the manuscript incorrectly referred to the intervals shown in Figure 4 as confidence intervals, although they corresponded to prediction intervals. Following the reviewer's recommendation and to ensure direct comparability with Cao and Furuya 2025, we have revised the analysis and now estimate the 95% uncertainty band of the fitted functions as the predicted value ± 1.96 times the residual standard error. This approach yields a constant-width band, consistent with the approach of Cao and Furuya 2025. The revised uncertainty bands are now shown in Figure 4 and the method is described in the newly added Section 4.5. For Figure 5, which shows the averaged trend line across all regions, we additionally provide a bootstrap confidence interval for comparison purposes.
- We applied the following changes to the manuscript:
- New line 503:
    - The 95% bootstrap confidence interval, incorporating all data points, suggests that alignment may occur as early as 43 years post-fire, whereas no alignment with the reference level is reached within the observed time period for the lower bound of the interval. The 95% uncertainty band based on the residual standard error, representing the variability of the observed data around the fitted function, suggests that alignment with the reference level could occur as early as 12 years post-fire.
- New line 508:
    - These findings differ from previous results on thaw season subsidence (Michaelides et al., 2019; Cao and Furuya, 2025) “based on the bootstrap confidence interval. However, the residual standard error based uncertainty band suggests that some individual fire scars may nevertheless fall within the recovery time ranges reported by previous studies.”
- Section 6.1:
    - Furthermore, uncertainties of the fitted functions and observations need to be considered. The uncertainty band based on the residual standard error represents the scatter of the observed data around the fitted function and reflects the variability among individual observations. Consequently, the broad uncertainty bands indicate that recovery times of individual fire scars may deviate from the recovery time inferred from the fitted function. In contrast, the bootstrap confidence interval reflects the uncertainty of the fitted recovery function itself.
- Section 6.3:
    - ..., which is much later than what is observed from the “fitted function of” the annual deformation rate values of our study, ... “However, the estimate reported by Michaelides et al. (2019) falls within the RSE-based uncertainty band of our fitted

function for this region and is also consistent with the uncertainty of the individual InSAR-derived deformation rate estimates. This suggests that, although our fitted function indicates a faster recovery, the associated uncertainties overlap with previous observations (Figure A8b)". Compared to Michaelides et al. (2019), our study incorporated a larger amount of fire scars and covered a longer temporal span of fire event years, which enabled a more comprehensive reconstruction of the system's temporal evolution "and may contribute to differences observed between the studies."

-I have a clarifying question on the method used for the unburned reference using a 5-km bufer: When you say "Areas with multiple fire events were excluded", does this mean that all locations afected by repeat burn were excluded? If yes (which I agree is important to remove), do the authors also account for any and all burn scars that fall within the 5-km bufer of nearby fire scars? Looking at the density of fire scars particularly in the YKD and the Noatak, there are several fire scars for which adjacent single-burn fire scars will fall within the 5-km bufer, but the text is not clear whether these pixels are excluded from the unburned bufer calculation. To be clear, I think that these pixels should be removed from the mean unburned reference estimate so as not to bias the 'burned-unburned' values reported in the manuscript.

- Reply: We confirm that all such areas were exclude. To clarify this we added the following to line 215:
    - … , which was defined as a 5-km buffer around the scar, "excluding all recorded fire scars within the buffer."

-On the choice of reference pixels: The authors' selection criteria for reference pixels is sensible and consistent with prior studies. However, I wonder what efect there is by choosing the airstrip in Bethel as a reference pixel for the YKD, when this is approximately 80 km from the study region, significantly farther than the 20 km radius used for the atmospheric high pass filter. It seems plausible that this choice could result in a residual low frequency phase term that is not accounted for in the high pass filter. Investigation of areas of barren ground within the study region, which one would expect to exhibit zero or near-zero annual deformation rates based on the authors' rationale for reference points in other study regions, instead exhibit large uplift rates of almost 3 cm/yr (see my very hastily constructed zoomed in screenshot with a few barren areas as identified from the Alaska NLCD 2016 circled in magenta):



The pervasive uplift signal to the southwest of the Yukon river, which corresponds to an area of boreal and mixed forest transitioning into low shrubland, also raises some concerns, and does not seem physically plausible. I suspect that if the authors instead used this barren ground closer to the study region for phase referencing, this would result in a systematic reduction in the annual deformation rate across the scene of about 2-3 cm/yr. I don't think that this would necessarily afect the 'burned-unburned' calculation that the authors use, although it would be worthwhile to confirm this. Nonetheless, this would certainly afect the annual deformation rate estimate. To the broader point of reference point selection – which is a notoriously dificult aspect of InSAR analysis in periglacial landscapes – it might be valuable to include some sort of assessment of the chosen reference pixels. One possibility would be to verify that the annual deformation rate over other pixels of barren ground also exhibit no significant deformation rate.

- Reply: We revised the Yukon annual deformation analysis by excluding the seasonal interferograms, as suggested by the reviewer. This eliminated the apparent uplift signal southwest of the Yukon River. Regarding the suggested barren areas, we note that these locations are situated within the floodplain/wetland complex of the Yukon River. Based on optical satellite imagery, these surfaces do not appear to represent stable barren ground. Instead, the CALU land-cover dataset classifies them predominantly as moist tundra with abundant forbs. Furthermore, we excluded areas adjacent to river channels in some of our other study regions because they frequently exhibit significantly different InSAR signals and, in this case, are located within an area surrounded by low-coherence. Therefore, we do not consider these locations suitable reference areas. In addition, the suggested barren areas are not equivalent to the bare ground reference areas used elsewhere in the manuscript, which consist of low-lying bedrock surfaces.
  Nevertheless, following the reviewer's suggestion, we tested the influence of using this alternative reference area on our results. We found no differences in the reported burned–unburned values, as this approach was specifically chosen, as mentioned in the manuscript, to mitigate, among others, the influence of potentially poorly selected reference points (see line 216).

-On the comparison with InSAR interannual deformation rates with Lidar ΔDTMs, line 422: "Moreover, the spatial distribution of pronounced annual subsidence rates observed in the DTMs in parts aligns with annual subsidence rate patterns derived from PALSAR-2, as illustrated in Figure 9. This similarity in spatial patterns is also reflected in a linear trend visible in the comparison of the value distributions 425 (Figure 10b)." I'm sorry, but I remain unconvinced that there is a robust, quantitative agreement between these two datasets. While the statistical distribution of values are broadly similar (Figure 10a), the violin plot (10b) demonstrates that there is both an additive bias (evident from the ofset between the 1-1 line and the line that would pass through the means of the violin plots) and a conditional bias (evident from the shallower slope of the line passing through the means of the violin plots compared to the 1-1 line) between the two estimates of annual deformation rate. The authors note that the two datasets broadly qualitatively agree 'in parts', but there are conversely many other places where the two datasets deviate rather significantly 'in parts'. For example, the authors note the pronounced annual subsidence rates observed in the DTMs align with subsidence observed from PALSAR-2, but there are also clear areas for which the DTMS record uplift while PALSAR-2 observes subsidence. If the authors intend to use the lidar ΔDTMs as a form of validation for the InSAR results, then this section would benefit from a more rigorous quantitative treatment of the biases between these two datasets and a discussion of whether a thorough quantitative validation is possible amidst such biases.

- Reply: In order to address these concerns, we have added a more quantitative treatment of the biases between the two datasets, adapted Figure 10b to better emphasize the identified biases, and added a Bland-Altman plot (Figure 10c) to further investigate the agreement between the datasets. We have made the following changes to the manuscript:

- Section title 4.4:
- Old:
  - Validation and qualitative assessment of InSAR results
- New:
  - Qualitative assessment of InSAR results
- Section title 5.6:
- Old:
  - Validation of InSAR annual deformation rate results
- New:
  - Qualitative assessment of InSAR annual deformation rate results
- New line 438:
- Old:
  - To assess the reliability of ...
- New:
  - To assess the spatial consistency of ...
- Extended section 5.6:
  - However, a quantitative comparison reveals systematic differences between the two datasets. An additive bias is indicated by the regression intercept of 0.35 cm/yr, while the regression slope of 0.10, which deviates substantially from the 1:1 line, indicates a conditional bias between the LiDAR-derived and InSAR-derived annual deformation rates. A Bland–Altman analysis (Figure 10c) further assesses the agreement between the two datasets, revealing a mean bias of −0.25 cm/yr and 95% limits of agreement ranging from −2.07 to +1.57 cm/yr, with most pixels falling within these limits. however, a proportional bias is evident, with differences increasing systematically for pixels with larger deformation rates.
- Discussion section (new line 474 ff):
- Old:
  - The validity of annual deformation rates was assessed by comparing our InSAR results with DTMs, demonstrating relative agreement in both median values and spatial deformation patterns, further indicated by linear trends in the value distributions (Figures 9 and 10).
- New:
  - The qualitative consistency of annual deformation rates was assessed by comparing our InSAR results with LiDAR-derived ΔDTMs, demonstrating similar median values and broad correspondence in spatial deformation patterns (Figures 9 and 10). However, systematic biases between the datasets were identified, which may be attributed to the different observation periods covered and fundamental differences in measuring principles between the two approaches.

References:

Liu, L., K. Schaefer, T. Zhang, and J. Wahr (2012), Estimating 1992–2000 average active layer thickness on the Alaskan North Slope from remotely sensed surface subsidence, *J. Geophys. Res.*, 117, F01005, doi:10.1029/2011JF002041.