

Thank you very much for the detailed review that helped improve the manuscript. We have addressed all comments below.

The authors' responses are written in blue font.

Lines 15-17: Can authors mention the type and density of vegetation present in the area? It is also an important parameter influencing the accuracy of the measurements and can be correlated to seasonal changes.

Thank you for the comment. You are correct in stating that seasonal changes in vegetation cover can also affect the correlation. It is, however, expected that large changes in the surface appearance due to vegetation would cause the correlation to fail, leading to missing pixels rather than a systematic bias pattern as produced by illumination changes. Our study area is located at high elevations (~2500 - 4500 m asl) and characterized by sparse vegetation cover, mostly shrub and grassland steppe (Angillieri et al., 2020). It is located in the transition zone from the humid foreland to the arid Central Andean Plateau. Seasonal greening is only observed in the lower part of the catchment below 3000 m that receives more moisture, while the upper hillslopes, including the landslide surface at 4000 m, are vegetation-free. The absence of dense vegetation makes the Del Medio catchment ideal for studying the impact of illumination changes, without vegetation effects. If vegetation were to compromise our results, we would expect to see an East-West gradient in the bias distribution, which, however, is not the case (the eastern part of the catchment receives more moisture and has a higher vegetation cover than the higher-elevation, western part). Instead, the measurement bias correlates well with moving topographic shadows.

We have added a sentence in the description of the study area to inform the reader about the local vegetation cover and added a paragraph in the results to explain why the effects of vegetation cover are disregarded:

“The influence of other seasonal variables, such as the changes in humidity and vegetation cover, is disregarded due to the scarcity of vegetation at 4000 m, particularly in the upper part of the Del Medio catchment. We do not observe an East-West gradient of the signal strength that mimics moisture transport and vegetation-cover gradients in the catchment (the lower elevation eastern part of the catchment receives more moisture and has a higher vegetation cover). Dense deciduous vegetation may compromise correlation results of cross-seasonal correlation pairs in other study areas. However, it is expected that changing vegetation cover results in random mismatches as the surface becomes increasingly dissimilar, and not a systematic bias as is the case for illumination changes.”

Lines 38-39: Why not mentioning the seasonal variations such as illumination, humidity, and vegetation changes that can create ground displacement measurement biases? Indeed seasonal shadow refers to exactly? Shadows occur because of topography or obstruction,

independently from the seasonal conditions; however their pattern is correlated to illumination conditions subject to seasonal changes (then at the origin of the oscillatory pattern in the time series). Please be more specific in the description of this process.

We have refined that sentence to include seasonal variations and make a distinction between shadows and seasonal illumination changes.

Lines 40-41: The sentence “Erroneous measurements will degrade the quality of inverted displacement time series, depending on their magnitude and distribution” reads odd. Erroneous measurements will anyways degrade the inverted displacements. It is the bias in the inverted displacements that will depend on the magnitude and distribution of these erroneous measurements. It is maybe just a matter of rephrasing.

We have removed the sentence to avoid confusion. The problem of erroneous measurements should be clear also from the previous lines.

Lines 47-48: What about humidity and vegetation changes?

See previous point. Vegetation changes are disregarded (a) due to the sparse cover and (b) because of the observed pattern, which suggests a strong link between measurement bias and topography, not to vegetation.

Lines 55-56: This is a great point.

Thank you.

Line 77: It would have been interesting to look at the vertical component from stereo imagery, at least for the cumulative displacement, to be able to infer the possible vertical deformation and rotation of the landslide body as a result of its downslope movement. I know it is not the main focus of this study though.

Agreed, but unfortunately, we did not have high-resolution multi-temporal stereo imagery available, and we also consider this to be outside the scope of this study.

Line 83: The displacement rate was indicated to be ~2.5 m/yr in the introduction. It is indicated 2-5 m/yr here. Which one is the correct one?

Changed to 2-5 m/yr. This referred to the previous paper (Mueting and Bookhagen, 2024) where only pairwise measurements were taken. In the coherent timeseries (after time-series inversion), the rates are, however, closer to 2.5 m/yr, thus the confusion.



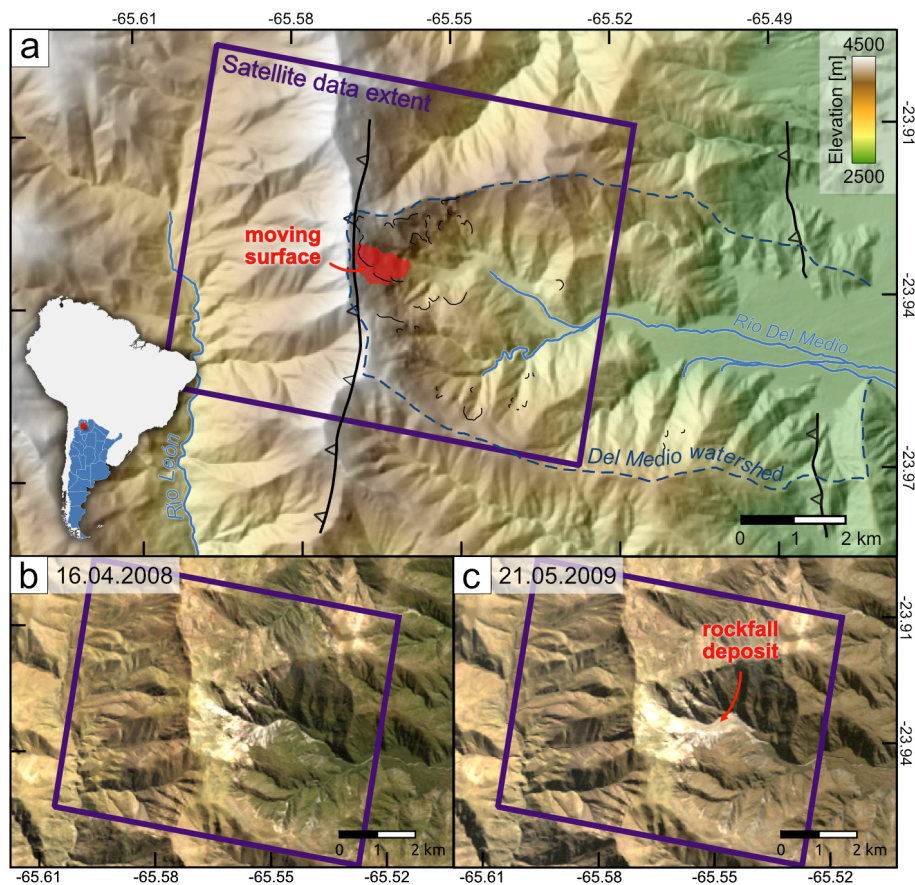
Lines 89-91: I agree that the question of whether the landslide behavior is being affected by the rainfall is an important question. Though the relation between rainfall and landslide have been well studied from observations on other case studies as well as analogue experiments. Authors should raise more specific questions that can be answered using the dense observations they propose, such as: Do the landslide movement reacts seasonally in response to the varying seasonal conditions? If not, what type of response does it show (linear for example) and why (controls of the kinematic of the sliding or bulk creeping)? Can we identify a delay between the start of the rainfall events and the landslide acceleration? Is there a minimum rainfall intensity (or volume) that is required before we start seeing any acceleration in the landslide movement? Further reading the study, it appears that the authors do not aim at studying any seasonal changes in the landslide behavior as it would appear correlated to the seasonal bias that is removed from the inverted time series (and only the inter-annual changes are kept). Authors should be more specific regarding the frequency of the landslide displacement signal they are trying to retrieve (and separate from the seasonal one). This is an unclear point all along the study, that I also raise at other places later in my review.

Figure 1: Please indicate the source of the DEM (SRTM, Copernicus?) and its ground resolution? Black lines showing the scarps and faults should be thicker, and possibly indicated in a legend directly in the figure. Text within panel (a) could also be larger (elevation legend, “moving surface”, and “satellite displacement”). Does the purple bow in (a) is similar to that in (b) and (c)? It seems like the ratio of the landslide size to the box size is bigger in (b) and (c). It could be nice to have an arrow indicated the landslide in (b) and (c). Eventually, can authors mention why the landslide is of a whitish color? Is it because of the lack of growing vegetation and soil on the surface that therefore reflects more light? Surprisingly, it appears more whitish in the “before” image, which could be related to the changing illumination conditions. To this regard, it would be interesting to have information on the sun elevation and azimuth to visually assess the relation between illumination conditions and shades in the images (as we see there are more shades in panel (c)).

The DEM is the Copernicus DEM (information added to the Figure caption). Lines and text size were enlarged. The purple box in (a), (b), and (c) is identical. Perhaps the ratio looked different because the linewidth was different, which has been adjusted. An arrow has been added to Panel (c) to indicate the location of the rockfall deposit. The landslide's brighter color is likely linked to the active erosion of the quartzitic rocks that are present in the upper part of the Del Medio watershed, while the vast majority of the catchment is dominated by darker Precambrian phyllites (see Figure 3a in Savi et al., 2016). However, illumination differences also affect the appearance because one scene is from October and the other from May, leading to the September scene appearing much brighter than the post-event image. We have therefore chosen to replace the October scene with a pre-event image that is more comparable in terms of illumination (~1 year difference) and appears less bright. Illumination conditions of all scenes are compared in the table below:

| Date             | Sun azimuth | Sun elevation |
|------------------|-------------|---------------|
| 2008-10-09       | 64.33       | 54.96         |
| 2009-05-21       | 38.39       | 34.92         |
| 2008-04-16 (new) | 46.83       | 42.31         |

This is the updated Figure:



Lines 101-102: What sampling strategy (frequency? Acquisition geometry? Illumination conditions?) was chosen to select the scenes when too many are available?

As mentioned in the text, imagery was chosen based on snow- and cloud-free conditions and AOI coverage. Illumination was not taken into account as a parameter (only later as a mitigation strategy), because we aimed to cover a wide range of illumination conditions. Acquisition geometry was only considered for PlanetScope scenes to minimize the effect of orthorectification errors. This pairing strategy is mentioned in the following paragraph on image correlation (Section 3.2). In case too many are available, a random scene was selected to have an acquisition approx. every 1-2 months. We have added a small clarification to the text.

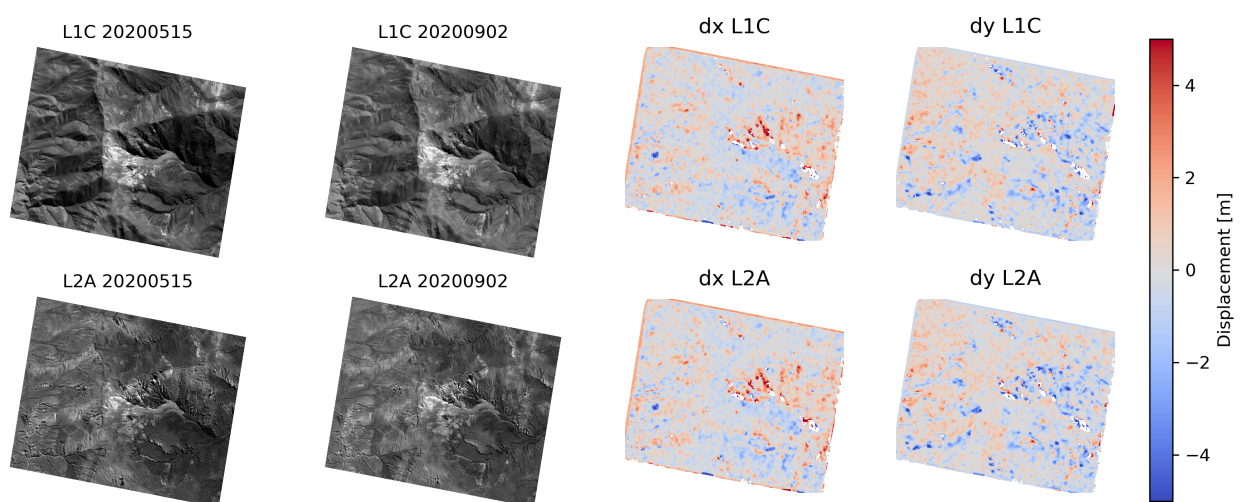
Line 103: I do understand that results from Lacroix et al (2018) show no difference in the results between the difference bands (or too small compared to the overall landslide displacement), though it would be interesting to test it here since the study specifically focuses on the effect of seasonal effects on the displacement measurements. The green band is the most sensitive to moisture changes. Sentinel-2 has a NIR band (band 8) that usually have a higher radiometric contrast, better feature stability, and less atmospheric noise. Even though all bands are correlated and that the results will not strongly be affected, it would be interesting to see how the choice of the band affects the results. See <https://custom-scripts.sentinel-hub.com/custom-scripts/sentinel-2/bands/> for a description of the characteristics of all the bands. On the same subject, it would be interesting to mention that one should not cross-correlate a reconstructed RGB image as each band have different geometric distortions which would add up in the reconstructed RGB and so the final correlation product.

We fully agree that a reconstructed RGB image should not be correlated since there may be interband misalignment issues, as e.g., documented for PlanetScope data (Aati et al., 2022). We have added a sentence to highlight this consideration.

A full comparison of the usage of different bands for all four different sensors used in this study is, however, out of scope and covered by previous work (e.g., Lacroix et al., 2018).

Line 105: Why not using the Sentinel2 L2A that have been corrected from atmospheric effects?

For Sentinel-2, L1C data is commonly used across many studies focusing on offset tracking applications (e.g., Lacroix et al., 2018). We made a comparison between the L1C and L2A products for an exemplary Sentinel-2 image pair, and the differences were very subtle (see below).



Line 117: Would it be possible to have an assessment of the effect of resampling (and so of the introduced aliasing) on the correlation results? It would be interesting to see how it affects the geometry of the shadows and of the resulting the displacement bias. Though I agree that upsampling is a better strategy than downsampling for preserving the PlanetScope image resolution and reducing the correlation noise. After further reading, I understand that the goal of upsampling the images in the current study is not to be able to cross-correlate images from the different sensors, but to compare correlation results made at similar ground resolution because of the effect of image resolution on noise (as proposed in Antoine and Liu (2025). Please specify this aspect as, from the current text, the reader is later expecting multi-sensor cross-correlations.

We have clarified in the text that only same-sensor image pairs are correlated, and the upsampling is done to better compare the resulting displacement fields, not the initial imagery. The effect of resampling is shown in Figure 5.

Line 125: Which half-year threshold? Maybe I missed this information earlier in the text.

This is an arbitrary threshold that was set to get an initial assessment of the expected displacement magnitude. It is not the final one. Based on the initial displacement estimate, we inferred which image pairs would be “worth” correlating, considering the image resolution and the detection threshold, assumed to be  $1/10^{\text{th}}$  of the native pixel size. This dynamic threshold results in more correlation pairs in times of rapid movement, while in times of relative quiescence, especially before the re-activation in early 2017, the temporal baselines are longer, and we do not expect to see any landslide displacement. This effect is well illustrated in the network plots in the supplementary material (Figures S3 – S5). We have also revised our text to better explain our dynamic thresholding approach.

Line 127: What is the range chosen for the common perspective ( $\pm 10^\circ$  around a given looking geometry for example)?

Much less than  $\pm 10^\circ$  – such a difference would introduce severe orthorectification errors given the large vertical changes that occurred at the landslide. The threshold set was  $\pm 0.6^\circ$ , analogous to Mueting et al. (2024). Most scenes, however, have a viewing angle difference less than that. We have clarified the threshold in the text.

Lines 128-130: It is maybe worth clarifying here that the study aims at correlating separately the different sates of images (L8, S2, and PlanetScope) to create three sets of displacement products that can be compared. Correlating all the pairs together is possible (if one focuses on the physical signal and needs as many data as possible); however, this would lead to mixing numerous confounding factors and a more difficult separation of the source different measurement biases (DEM, image and sensor distortions etc).

We have added a clarification that the matching across sensors is not advised, because it introduces further complexities.

Line 133: Do the authors considered a uniform shift over the entire image or was polynomial form fit to the correlation results (with the landslides masked)? Can the authors justify the choice of a uniform shift. Sentinel2 and Landsat8 images are not subject to the same type or image distortions and orthorectification biases than higher-resolution images acquired with more oblique geometries; however, I am wondering whether this uniform assumption is entirely valid.

All postprocessing modifies the original measurements to some degree. When there is no need for more complex geometric corrections, a simple shift is the most rigid transformation that can be applied. Since we did not observe any large-scale ramp or even more complex errors (as is the case for PlanetScope) in the Sentinel-2 and Landsat-8 displacement fields, we believe our choice of a uniform shift is justified and helps to preserve the original measurements as closely as possible. Also note that our study area is small compared to a full Sentinel-2 or Landsat-8 scene. At a larger scale, perhaps a polynomial fit correction would be more appropriate.

Line 136-137: How is the topography accounted for in the polynomial correction. This is not clear. Please elaborate as it is an important point.

The polynomial is fit as follows:

$$dx, dy = aX^2 + bY^2 + cZ^2 + dXY + eXZ + fYZ + gX + hY + iZ + j$$

where  $dx$  and  $dy$  are the estimated displacements in an EW and NS direction,  $X$  and  $Y$  are the corresponding pixel locations,  $Z$  is the elevation (derived from an external reference DEM), and  $a-j$  are the coefficients which are estimated using least-squares optimization (Mueiting et al., 2024).

We have modified the sentence in the manuscript to clarify the variables taken into account for the fit, but refer the reader to the relevant study for details.

Lines 156-164: This is a very interesting approach! However, I suppose that ASP gives a “correlation score” map at the end of the correlation process as COSI-Corr and MicMac do. In this case, why not using it to weight the different pixels as well? It could also be interesting to see if weighting the displacement maps based on the temporal difference between correlated pairs would have an effect on the inversion result (as in Lacroix et al., 2019). I see that temporal difference is accounted for in the regularization matrix when inverting the PlanetScope correlation results, but this is different and not applied to S2 and L8. I guess in the case of a more complex acquisition geometry, one could use the same principle to weight the displacement maps based on the similarity of image look angles (but this is not



the case of S2 and L8, and the authors already put a constraint on the PlanetScope image geometry).

Weighting based on the correlation score would indeed be an interesting approach to assess. However, a detailed quality metric is unfortunately not an output that ASP provides. The output is a 3-band raster containing displacements in the EW and NS directions, as well as a binary “Good pixel mask” which is applied to mask unreliable measurements. This mask typically only removes false matches in areas with large shadows or where the surface has become too dissimilar.

Line 186-187: It is not clear here which signal is preserved to analyze the landslide displacement. In fact, I guess the non-linear inter-annual displacement trend is not “removed” as it contains the landslide displacement signal. It is rather “separated” from the rest of the signal that likely include the seasonal variations that can be fit with a sinusoidal function with a frequency of 1 year, and a possible additional linear trend. In this case, this separated signal (using the 13-month kernel) will smooth out any higher-frequency variations of the landslide displacement. This suggest that the authors do not aim at studying the correlation between seasonal water load changes and landslide displacement, but more likely interannual changes due to longer-period climatic variations (El Nino/La Nina for example). Please rephrase, and add the mentioned specifications regarding the higher-frequency (maybe seasonal) landslide displacement signal that will be overlooked. I also suggest the authors add examples of uncorrected and corrected time series with the different components separated (as for example Figure 3 in Lacroix et al., 2019). This would significantly help the reader understand the post-processing step taken by the authors.

We have rephrased Section 3.5 describing estimation of seasonal biases, as it seemed the initial version had caused a misunderstanding. First of all, please note that this section describes our method to estimate the magnitude of the seasonal bias, not the correction per se, although it works alike. The trend (using the 13-month kernel) is only estimated and removed from the displacement signal to obtain the basis for the sine fitting (i.e., center the displacement at zero). Fitting a linear trend here is not useful because the landslide does not move uniformly over time. The smoothed trend will, as you mention, reflect interannual changes only, which is in fact the objective. When interannual changes are removed from the time series, the remaining signal reflects seasonal landslide changes + seasonal oscillations from the illumination bias. To the detrended time series, we fit a sine function whose amplitude is used to quantify the impact of seasonal bias.

Regarding the correction, this best-fit sine fit is subtracted from the original time series. The smoothed trend is not considered further. Therefore, we can assess seasonal velocity changes even after correction. The assumption here is that, unlike bias from illumination changes, seasonally variable landslide displacement (a) only has one direction (downhill) and (b) may be of variable magnitude, or not observed every rainy season. We have evaluated several synthetic displacement signals, to which degree this correction approach is viable in the presence of a true seasonal displacement component (see our response to this specific matter below.

There is a Figure in the Supplementary Material that illustrates the sine-fitting approach (Figure S8 in the revised manuscript).

Figure 2: In relation with my previous comment, it is not very clear whereas the time series of displacement presented in panels d-f were corrected from the seasonal variations. From the text, it seems like this figure presents the final product (after correction) though the seasonal variations are really visible. Are these residuals after correction? Or maybe the un-corrected time series are presented at this stage?

Figure 2 shows the uncorrected time series, while the corrected time series is shown in Figure 7. We have modified the text to make this clearer.

Line 227: Are the seasonal landslide displacement variations really analyzed or only the inter-annual ones? Please refer to my previous comment regarding the post-processing strategy and the separation of the different components.

It is correct that we look for seasonal landslide variations (and inter-annual ones as well). Please see our response to the previous comments.

Figure 3: Which period of the modeled seasonal signal is currently shown? Please specify. Moreover, there is a seasonal East-West signal at the location of the landslide. How do the authors treat this signal given that it could both be measurement bias but also seasonal changes in the landslide behavior triggered by temperature and humidity changes? Please elaborate on this aspect.

There is only one amplitude of the sine function, which is constant for all periods and plotted in Figure 3. And yes, the separation between illumination bias and potential true responses in landslide motion is the objective of employing different correction methods. This problem is discussed in the dedicated Section 4.4.

Line 236: This is because decreasing the pixel size and therefore the correlation window size will reduce the introduction of larger scale and amplitude seasonal signal. Maybe specify.

But the effect was also observed when the correlation kernel was kept approximately constant in terms of actual area covered, making sure that the same features are included. Please see the adapted kernel sizes depending on the input resolution in Figure 5, and also our response to the reviewer's previous point on line 243.

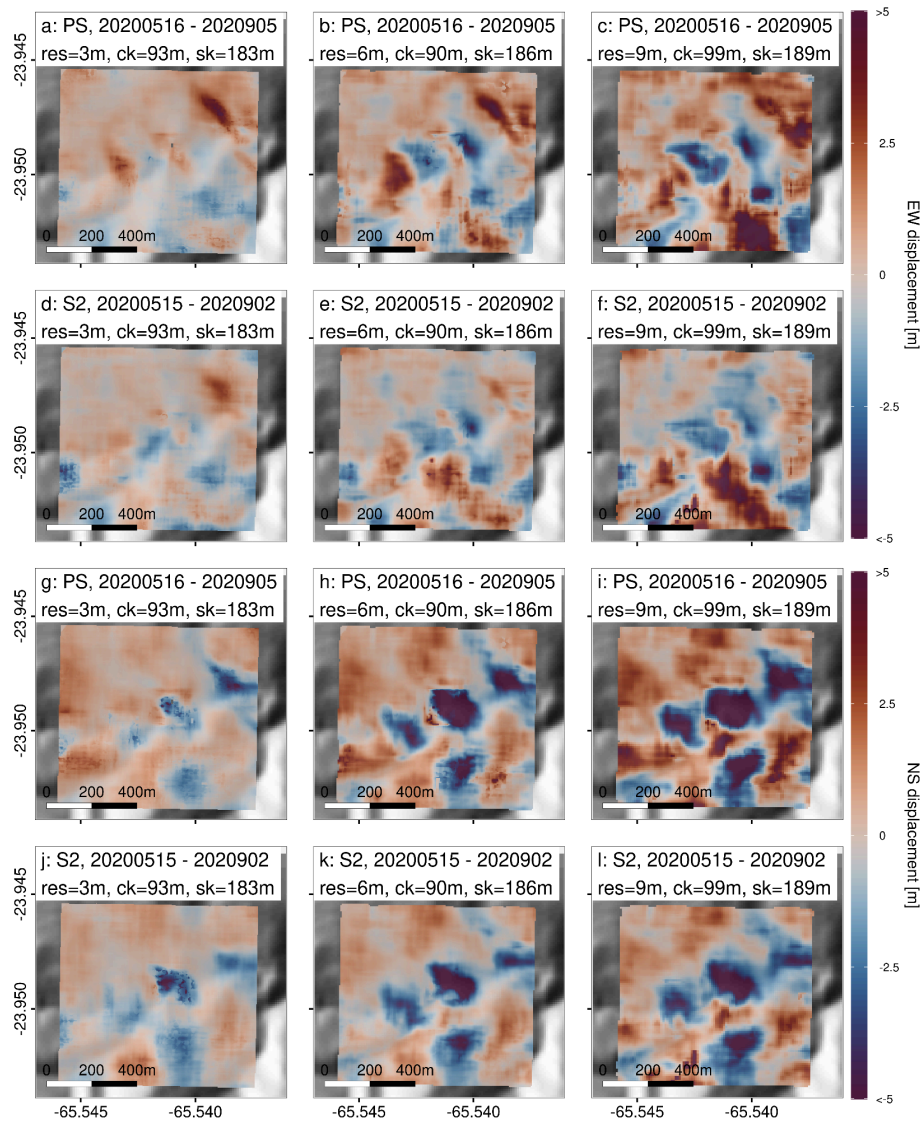
Line 241: The term “however” suggests a difference between the method used and the other ones. Maybe use “also” as the effect of resolution is seen for all methods?

“However” was exchanged with “also”.

Line 242: Of course, the subpixel matching precision depends on the resolution as it is usually around  $1/10^{\text{th}}$  of a pixel. However, I am not sure that the matching between upsampled pixels will be more precise as there is redundancy in the information and a greater number of positively matching pixels within a given area. This would need to be studied in more detail.

Indeed, the upsampling does not add additional information. For detecting finer-scale spatial variability in landslide motion, the use of higher-resolution data is therefore essential. However, for comparability across sensors, resampling to a common resolution is useful to ensure that results are produced using common processing parameters (e.g., same kernel size, same number of pixels used for matching). We have updated our analysis comparing Sentinel-2 and PlanetScope resampled to different resolutions to not only include a downsampling, but also an upsampling example. We used images from similar dates over the same area and observed similar effects, with the Sentinel-2 displacement fields (lower native resolution) appearing slightly smoother. The results for Sentinel-2 are now included in Figure 5:



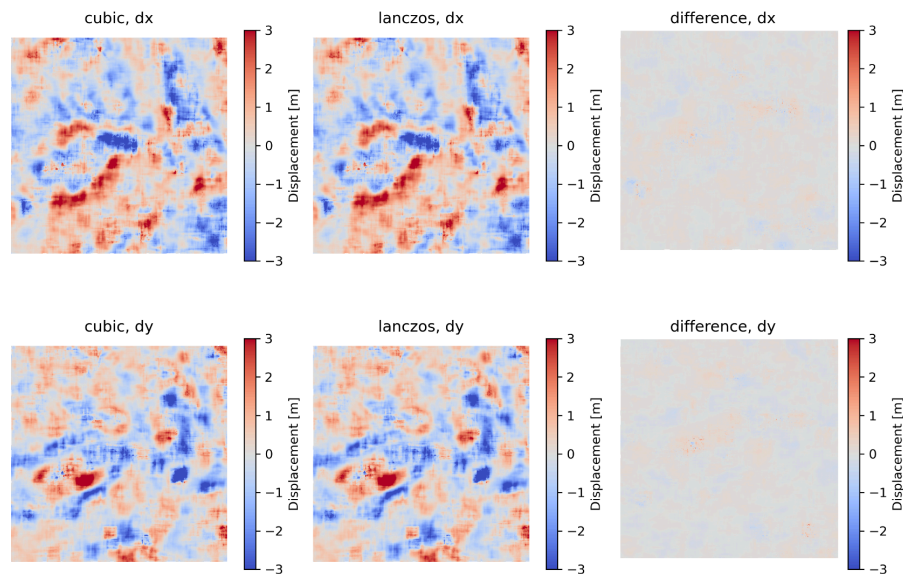


Line 243: Why would the number of pixels be different in the correlation windows if the correlation windows are set in pixel units?

See our response to one of the previous points. The correlation kernel was kept approximately constant in terms of metric area covered, which means fewer pixels at lower spatial resolution.

Line 244: I do agree that the amplitude of the seasonal signal that will be captured is function of the image resolution (as discussed right before), and therefore that it makes sense to correlate the images at similar resolutions if one wants to focus on the seasonal signal. however, there will be aliasing introduced into the correlation results, affecting the subsequent time series. How is this identified?

We did not observe any strong aliasing effects. We did a test repeating the correlation with a clip of two Sentinel-2 scenes over stable terrain (20200515 and 20200902). We upsampled these clips once with cubic interpolation and once with Lanczos resampling, which reduces aliasing artifacts. The displacement fields and the difference between them are shown below. We consider the difference minor, especially in comparison with the magnitude of biases introduced by the variable illumination.



Line 255: Missing “,” between “For the correction” and “Lacroix et al. (2019)”.

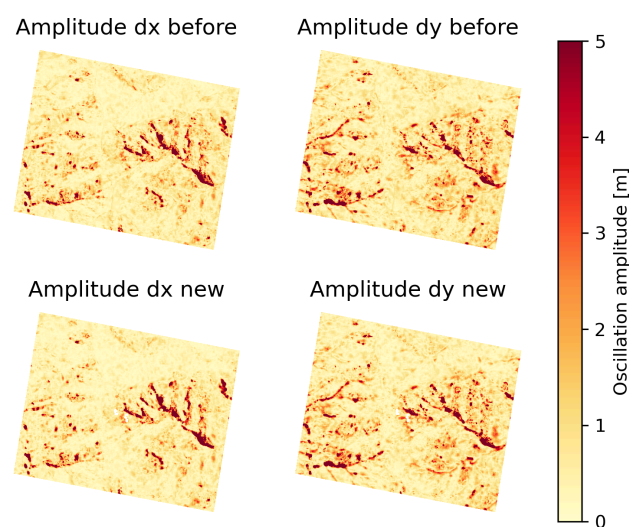
A comma was added.

Lines 265-268: Could a high-resolution DEM be derived from Pleiades stereo imagery for that area to resolve this issue? Maybe even multiple Pleiades DEMs could be derived if more imagery is available. WorldView imagery could also be used but accessing those images is more complicated unless authors have an ongoing collaboration or project with American institutions such as NASA and USGS.

Yes, as mentioned in the text, we do suspect that a higher-resolution DEM depicting topography closer in time could improve this correction approach. We do, however, believe that it is of interest to evaluate how well the mitigation can be performed using more easily accessible and global elevation data, such as the Copernicus DEM.

Lines 272-275: Would it be possible to assess a per-pixel oscillation in stable terrain areas and interpolate over the landslide area, therefore not including the true landslide displacement? This would not be ideal as slip and aspect can be different at the location of the landslide compared to its surrounding, but I am wondering whether it could be an option in some cases.

Yes, this is the approach that was described by Lacroix et al. (2019). We tried to replicate that approach and estimated the seasonal bias signal for stable areas using bins of common slope and aspect and extrapolated these values to the landslide surface (sine fit stable areas). However, we observed that the reduction was not as effective compared to a pure per-pixel approach, likely because of the lower resolution topography and topographic changes that occurred over the landslide since the DEM acquisition. In comparison with the initial version of the manuscript, we refined this approach slightly by updating the topographic groups using slope and aspect calculated on the initial DEM resolution. We also subtracted the median, not the mean, time series instead to reduce the influence of large errors from mismatches in heavily shaded regions. This slightly improved the results for this correction method, but not substantially, see below (also updated in Figure 6):



Line 77: Should it be “responses of the displacement” instead of “responses to the displacement”? What are the arguments for supporting that seasonal response should be variable? If the structure of the landslide has not changed much over one year, its response will also be the same isn’t?

Thank you for spotting this; it should indeed be “responses of the displacement”, or even better, “responses of the landslide motion”. A variable response could be linked to variable precipitation rates in different years – some rainy periods may be drier than others. For the landslide, we postulate that acceleration is linked to single extreme rainstorm events – these may not occur every year.

Line 279: What about other strategies based on ICA decompositions? These are proven quite efficient at separating signal from different source. I think these following studies used

ICA to decompose the signal: Ali et al. (2020), Beaud et al. (2022), and Bontemps et al. (2018).

According to our understanding, these studies did not use ICA decomposition to assess the impact of seasonality. Ali et al. (2020) used differences in solar illumination as a pairing criterion to reduce the impact of seasonal bias, Beaud et al. (2022) used principal component analysis to reduce the noise level, and Bontemps et al. (2018) evaluated seasonal biases as a source of uncertainty, but did not choose a dedicated mitigation approach.

We believe the reviewer may be referring to seasonal decomposition strategies into a trend, seasonal, and noise/residual component. The problem here is that a recurring seasonal signal in the landslide motion would be incorporated into the seasonal component. Our objective is to unmix true seasonality from false seasonality, i.e., illumination bias. We therefore approximate the seasonal bias with a sine wave so that anything deviating from that (e.g., directed response to landslide from rainfall, downhill only) is not removed from the time series.

Line 301-302: Very nice approach indeed! However, I do not understand why the temporal difference is not also accounted for in the image-pair selection. Lacroix et al. (2019) proposed an approach to do so.

We did take into account the temporal difference between scenes; however, we did not set a maximum threshold but a minimum temporal baseline to be able to detect landslide displacement. We did not restrict correlation pairs with large temporal differences because, as argued before, vegetation effects and anthropogenic influence, which would cause decorrelation over time, are limited in our study area.

Lines 303-311: This is a good point. But maybe the best strategy is then to keep all the pairs, but weighting the correlation products based on the temporal, illumination, and viewing angle similarities)

We did discuss and experiment with different weighting approaches. In this study, however, we focused on weighting factors that would address seasonal biases. But yes, to mitigate other effects, a weight based on multiple factors could be considered, e.g. temporal baseline, correlation score (when available), or sun-angle differences.

Line 310: I do not understand this part. One would expect seasonal biases to have different characteristics during different seasons. Moreover, the regular sampling mentioned has a period of 1-month. I do not understand how a seasonal signal would average to zero over a period of 1 month. Please clarify.

For this synthetic example, we introduced errors that were dependent on the temporal baseline between two dates (the highest error for scenes with a half-year difference). But

indeed, we did not consider whether specific season matches have more errors than others. Because the impact of temporal sampling is not further explored in the manuscript, we decided to remove this sentence and the corresponding Figure from the Supplementary Material to avoid confusion.

Lines 324-326: The weighting approach seems more efficient (similar to the sine correction) for the PlanetScope time series compared to the L8 and S2 ones. Maybe this could be due to the fact that L8 and S2 images are acquired overall in similar illumination and look angle conditions (at the same time in general) compared to the PlanetScope that has a more variable imaging schedule and geometry? Then the weighting would be more impactful. Eventually, there is a shift between the original and the corrected time series derived from L8 imagery. Does this come from additional corrections which are not mentioned here? Please elaborate.

In Figure 6, we only compared the different mitigation approaches for PlanetScope data, not for different sensors. In Figure 7, on the other hand, we compare the two most effective approaches (sine fit per pixel and pair selection based on similar illumination conditions, not weighting) for Landsat-8, Sentinel-2, and PlanetScope. In the revised manuscript, we have updated Figure 7 to include all correction approaches, even when there is very little difference to the original signal. In the initial Figure, indeed, a shift was observed for the Landsat-8 time series. We have investigated this, also concerning a point raised by the second reviewer, and found that this was linked to noisy pixels around the border of the landslide. In the revised manuscript, we have updated our landslide mask by applying an inward buffer of half the correlation kernel size, which excludes pixels close to the transition between stable and moving terrain. All applicable Figures were updated accordingly.

Line 360: The lack of correlation between the landslide movement and the seasonal variations has not been discussed much. This needs more elaborated discussion before assuming there is no strong correlation. See also my previous comment regarding the frequencies of the signal that are separated, and how some of the landslide seasonal signal could have been missed.

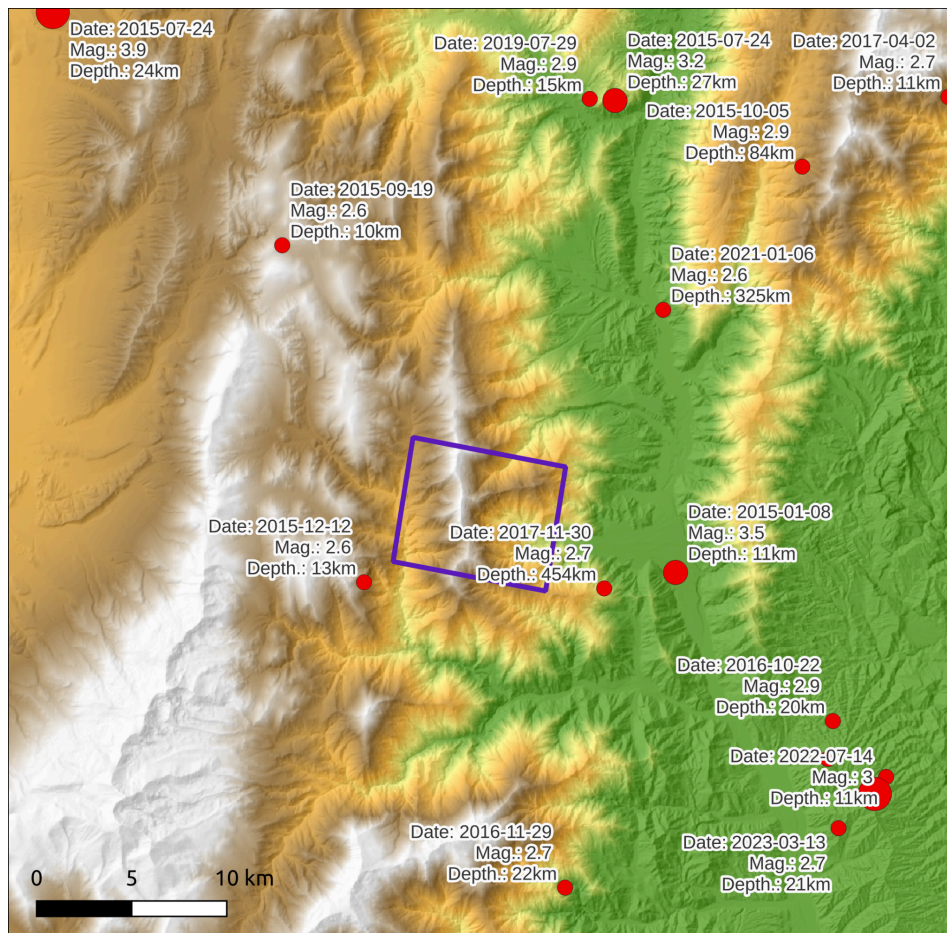
We have elaborated more on the observed signal and the separation between seasonal bias and true displacement, and supported this through further assessment using a synthetic displacement signal. It is true that by using a sine-fit per pixel, part of a true seasonal component could have been removed. Since no ground control data is available for the landslide, reducing seasonal biases through selecting similarly illuminated scene pairs is likely the safest choice at the expense of increasing the impact of errors linked to other sources due to the lower number of overall measurements. We have adapted our discussion to reflect this.

Line 367: A radius of 100 km is not enough if a large subduction earthquake; however, there is no record of such event in 2017 (Michel et al., 2023). Besides this, authors need to support



their statement citing a seismicity catalogue for the given area. All earthquakes are not published on the USGS website or other if they are not large or impactful enough. Maybe small earthquakes occurred locally?

Thank you for the remark. We did check an additional local earthquake catalog from the Instituto Nacional de Prevención Sísmica ([http://contenidos.inpres.gob.ar/buscar\\_sismo](http://contenidos.inpres.gob.ar/buscar_sismo)). Indeed, this included smaller earthquakes closer to the landslide; however, these events occurred either not on a date matching the acceleration phases, at great depth, or had a very low magnitude (see the map below with the study area marked by a purple rectangle). We have added a clarification to the text.



Line 381: This still remains to be better demonstrated (see my previous related comment).

We have supported this conclusion with additional tests using synthetic data and a more careful evaluation of the effects of different correction approaches.

Line 382: This sentence reads incorrectly. The “periods” are not “documented in landslide motion”. I guess what the authors meant to write is that studies showed that landslide motion



can be significantly influenced by periods of extreme draught and rainfall. Please correct the sentence.

The sentence has been corrected.

Figure 10: What is the reference used to calculate the difference maps? This is not indicated in the caption nor in the figure. Did the author do a difference between the landslide displacement map (corrected) and the uncorrected map, and so separately for each sensor? Can authors also indicate the location of the box chosen for plotting the time series in j-k? (see another comment below to this regard). This should also be indicated for profiles from Figures 2, 7 and 9.

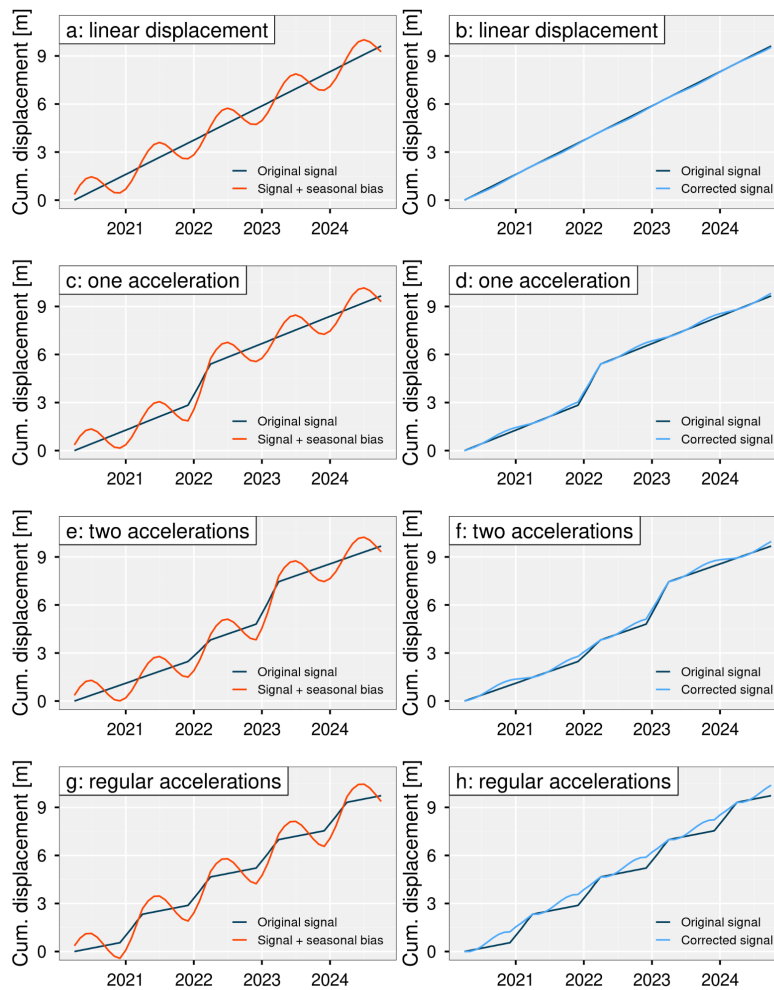
There is no fixed reference; the difference shown is always between the total displacement estimated for two different sensors: PlanetScope - Landsat-8 (a, d), PlanetScope - Sentinel-2 (b, e), and Sentinel-2 - Landsat-8 (c, f). This is indicated in the description of the individual panels. The first row shows the uncorrected displacements, the second row the displacements corrected through a sine fit per pixel. We have added the landslide outline used to extract the displacement time series in panels j-k.

429: I do not understand this sentence. The closest correspondence rather seems to be achieved between the S2 and PlanetScope images. The L8 time series analysis rather shows a deviation towards the later dates with a lower inferred cumulative displacement over the total period (Figure 10 j,k). Moreover, what could explain that less cumulative displacement is detected using L8 images? Can this be related to the image native resolution that is lower and therefore not capturing more local and higher amplitude displacements? Or maybe to a lower image radiometric or geometric (orthorectification) quality? Or maybe this is just a matter of spatial representation of the point chosen for displaying the time series. In fact, in Figure 8, it does not seem that the total displacement inferred from the L8 images is lower than inferred from the S2 and PlanetScope images. Can authors explain better how they select the point for displaying the time series (location, width, and interpolation if choosing a box)? The location and width of this area should also be indicated in all the figures.

We hope Figure 10 has become clearer with the updated Figure labels. With closest correspondence, we refer to the total estimated displacement over the entire terrain, not the landslide time series. For the pixel extraction over the landslide surface, we have updated the relevant mask to exclude border effects (see our related reply to an earlier comment). This mask was added to the maps shown in Panels a-f.

Line 458: Indeed, the correction method used can remove seasonal landslide displacement signal. This is not well discussed in the manuscript where authors assume several times that the landslide does not present seasonal motion, but maybe it has just been removed during post-processing. This is an important point of this manuscript that needs to be clarified.

We have updated the manuscript with additional analyses on what the effect of the sine fitting approach may have for different signals (linear, individual acceleration phases, regular seasonal responses). The results are shown in the Supplementary Material, Figure S21:



The per-pixel sine correction approach is most effective when the occurrence and magnitude of true seasonality are irregular. For the Del Medio landslide, our interpretation of the signal observed (across different correction approaches) points more towards irregular accelerations linked to extreme rainfall events. True seasonal responses would be expected to also reflect in the time series retrieved from similarly illuminated pairs. However, since there is no ground control data for the landslide, we conclude that the time series based on image pairs with similar illumination is likely the safest choice for the subsequent interpretation of the landslide signal, since this fully avoids any alteration of the initial measurements. Closely assessing the PlanetScope time series from similarly illuminated pairs only, we find that there may even be two more small acceleration phases that were not visible in the sine-corrected time series.

Eventually, all correction methods come with a trade-off and should be chosen based on the study area and available data. We have extended the discussion in the manuscript regarding the benefits and limitations of each mitigation strategy.

Line 468: Indeed, weighting factor based on temporal changes might be more efficient. Please mention it.

We have added this for consideration.

Line 469: Choosing this specific approach implies that authors focus on the inter-annual signal, and do not aim at analyzing the possible seasonal landslide movement variations (removed during the correction using the since function). Please specify.

Please see the earlier clarifications. The smoothed trend is just removed for the estimation of the sine fit, but the actual time series is not smoothed. That is why seasonal changes can still be detected, provided that they exceed a certain magnitude.

Line 479: This remains to be better demonstrated (see all my previous comments related to this point).

Please see the updated interpretation of the displacement time series based on image pairs with similar illumination and the related discussion in the revised manuscript (Sections 4.5 and 5.1).

Line 480: What about fire? Was there any significant fire that would have removed the vegetation and therefore making the landslide more unstable? Just searching in Google, I could find there were significant fires in 2016 and 2017 in Argentina in the provinces of Jujuy (where the study area is) and Salta (just south of it) that were affected. See <https://www.batimes.com.ar/news/argentina/argentina-ablaze.phtml>. See also interesting references in <https://science.nasa.gov/earth/earth-observatory/fire-threatens-rare-forests-in-argentina/>. I think this is worth looking into as there is a known correlation between fires and landslides (Kean and Staley, 2021; and <https://scienceexchange.caltech.edu/topics/sustainability/ask-expert-sustainability/debris-flows-michael-lamb>).

Thank you for this interesting remark. Although both references refer to debris flows rather than slow-moving landslides, we have checked satellite imagery of the Quebrada del Humahuaca from 2017 and 2018 and did not observe any burn scars. In addition, we have consulted the MODIS Fire\_cci Burned Area Pixel product (Padilla Parellada, 2018) for the years 2014-2024. This dataset indicates the large majority of burned areas further to the East in the Andean foreland beyond the drainage divide of the Humahuaca basin, where more dense vegetation (potential fuel) and agricultural areas can be found. No burned areas were detected in the Del Medio basin.

Lines 591-592: The link <https://doi.org/10.5194/egusphere-2023-1698> in “Mueting, A. and Bookhagen, B.: Tracking slow-moving landslides with PlanetScope data: new perspectives on the satellite’s perspective, EGU sphere, 2023, 1–36, <https://doi.org/10.5194/egusphere-2023-1698>, 2023.” Takes to the preprint, not the journal article. Please update.

Thank you for catching this. The reference has been updated.

## References:

Aati, S., Avouac, J.-P., Rupnik, E., and Deseilligny, M.-P.: Potential and Limitation of PlanetScope Images for 2-D and 3-D Earth Surface Monitoring With Example of Applications to Glaciers and Earthquakes, *IEEE T. Geosci. Remote*, 60, 1–19, <https://doi.org/10.1109/TGRS.2022.3215821>, 2022.

Angillieri, M. Y. E., Perucca, L., and Vargas, N.: Spatial and temporal analysis of debris flow occurrence in three adjacent basins of the western margin of Grande River: Quebrada de Humahuaca, Jujuy, Argentina, *Geografiska Annaler: Series A, Physical Geography*, 102, 83–103, <https://doi.org/10.1080/04353676.2020.1744075>, 2020.

Lacroix, P., Bièvre, G., Pathier, E., Kniess, U., and Jongmans, D.: Use of Sentinel-2 images for the detection of precursory motions before landslide failures, *Remote Sensing of Environment*, 215, 507–516, <https://doi.org/https://doi.org/10.1016/j.rse.2018.03.042>, 2018.

Mueting, A. and Bookhagen, B.: Tracking slow-moving landslides with PlanetScope data: new perspectives on the satellite's perspective, *Earth Surf. Dynam.*, 12, 1121–1143, <https://doi.org/10.5194/esurf-12-1121-2024>, 2024.

Padilla Parellada, M. (2018): ESA Fire Climate Change Initiative (Fire\_cci): MODIS Fire\_cci Burned Area Pixel product, version 5.1. Centre for Environmental Data Analysis, 01 November 2018. <https://doi.org/10.5285/58f00d8814064b79a0c49662ad3af537>.

Thank you very much for the detailed review that helped improve the manuscript. We have addressed all comments below.

The authors' responses are written in blue font.

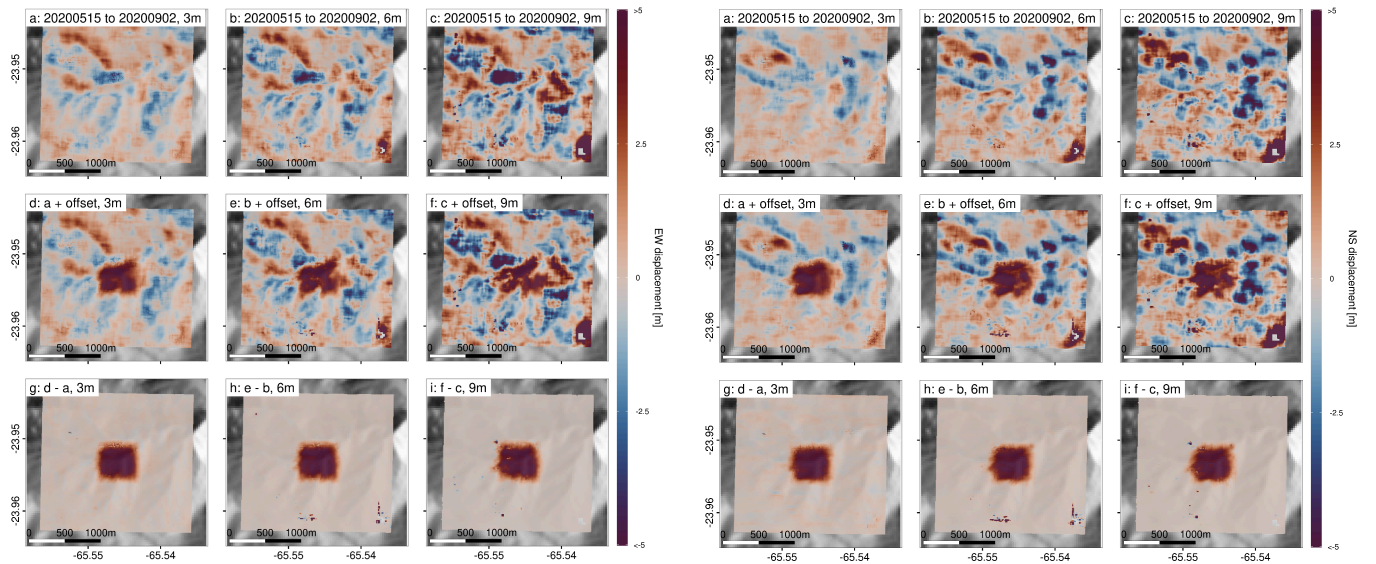
The comparison of the different mitigation strategies are valid on the stable areas, but are more questionable on moving areas. Because the correlation operator is not linear, I am wondering what is the combined effect of bias illumination variation and a moving target on the results. Since no validation datasets exist on the Del-Medio landslide, this question is still open.

We believe that the Del Medio landslide is a great case study to highlight the problem that seasonal biases pose for interpreting landslide motion in a seasonal climate. As the reviewer probably knows well, it is unfortunately very difficult to find multi-annual, publicly available GNSS data on slow-moving landslides that are well trackable with optical imagery (meter scale displacements, sparse vegetation, low cloud cover), let alone a landslide matching these criteria that is locally monitored and shows a seasonal displacement component.

While there is no GNSS data available for this landslide, there are different ways to distinguish between the true signal and a measurement bias: (1) over stable areas, the displacement is expected to be zero; (2) over the landslide surface, we expect mostly (only) downhill motion; (3) estimated displacement rates are (approximately) similar across all datasets used.

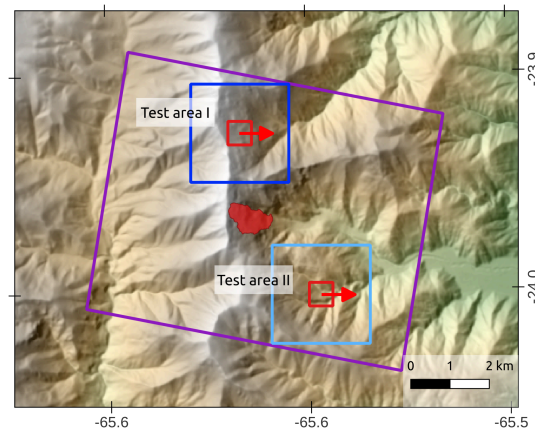
Regarding the combined effect of illumination bias and moving targets, we have added results of an experiment to the Supplementary Material, Figures S19 and S20 of the revised manuscript. We correlated two Sentinel-2 images from different seasons (2020-05-15 and 2020-09-02) over stable terrain with different illumination conditions (and also upsampled to test the effect at different resolutions). We then introduced an artificial and known offset to the secondary image (500x500 m patch, shifted by 5 m in both x and y directions in the center of the scene) and ran the correlation again. As expected, the seasonal illumination bias obscured the true signal. However, when the illumination bias is removed (by subtracting the displacement estimated without an artificial offset), the remaining displacement field clearly shows the moving patch and zero displacement estimated over stable terrain. This leads us to the conclusion that (a) with a perfect correction method, it is possible to fully recover the true displacement signal and (b) the success of correction over stable terrain can be transferred to the landslide as well. Furthermore, for our scenario, the bias introduced by the illumination is the dominant source of noise.





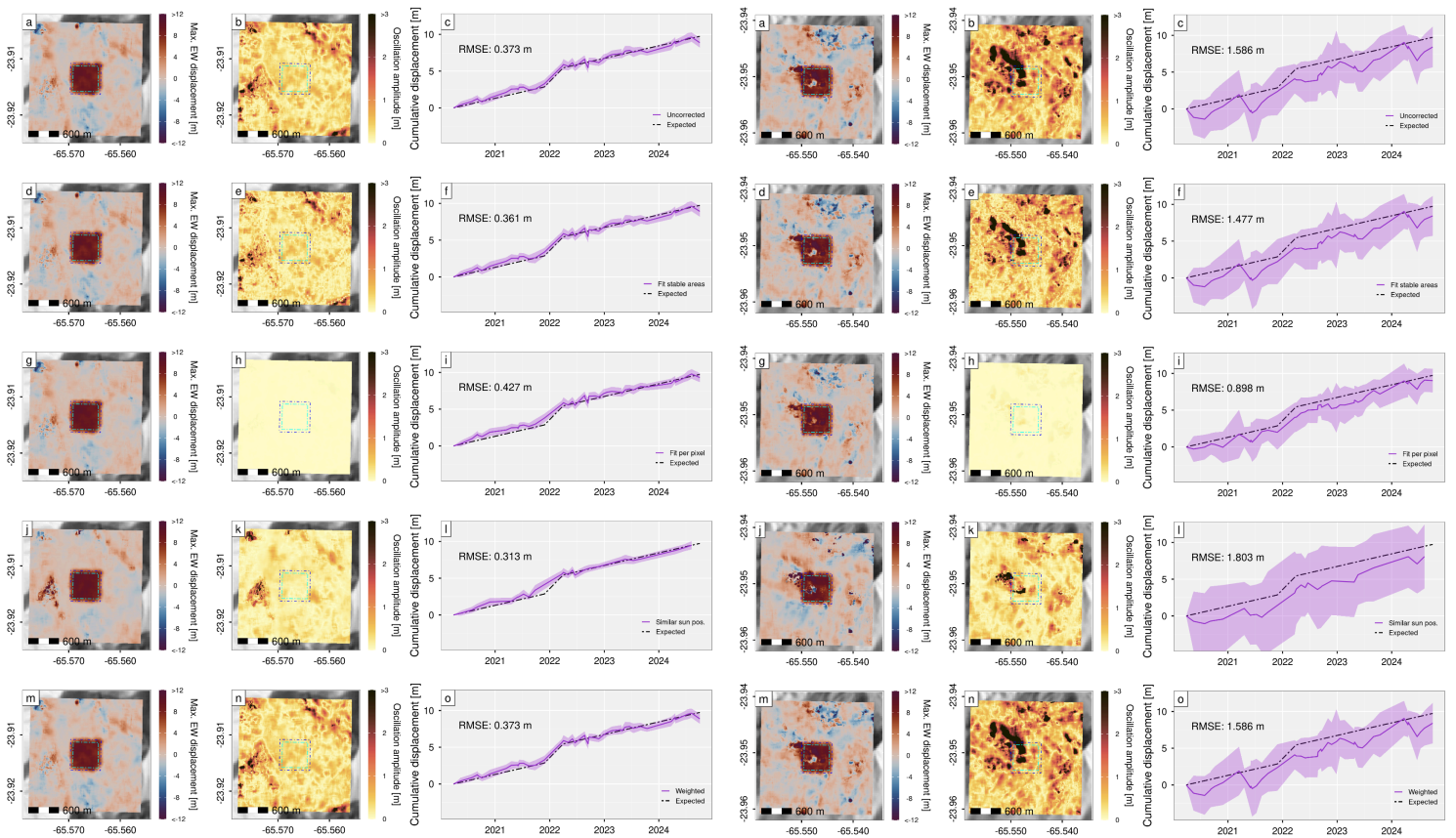
Figures S19-S20: Displacement fields estimated over stable terrain using a cross-seasonal Sentinel-2 image pair (May and September 2020), resampled to 3 m (first column), 6 m (second column), and 9 m (third column) spatial resolution. The first row shows the seasonal biases that affect the displacement estimates when the original scenes are correlated. To simulate the mixing of seasonal bias and true displacement, we introduced an artificial offset to the secondary image by moving a small patch in the center of the scene by 5 m in EW and NS directions. The results of the correlation, including the added displacement, are shown in the second row. When subtracting the pure seasonal bias (first row) from the mixed signal (second row), the true displacement can be reconstructed well (third row).

Introducing artificial displacement over stable terrain also opens up the opportunity to evaluate the full processing chain and the effect of seasonal bias mitigation. We did so using all available PlanetScope imagery from the newest generation over two stable areas: one well-illuminated area with a similar orientation as the landslide (test area I) and one heavily affected by shadows (test area II), see map on the right (Supplementary Figure S22 in the revised manuscript). In both areas, we moved a small patch by a known amount, approximately reflecting what was observed over the Del Medio landslide between 2020-2024 (linear displacement (total of 10 m)



with a small acceleration (2 m in 4 months) in the monsoon season of 2021/2022.

The results look as follows for test area I (left) and area II (right) (see previous Figure for the location). Plots show the maximum EW displacement (first column), oscillation amplitude (second column), and the mean displacement extracted over the known moving area (with an inward buffer to avoid border effect, see also our response to the comment regarding the border effects) (third column). The rows show the respective correction approaches: Uncorrected (row 1), sine fit over stable areas (row 2), sine fit per pixel (row 3), similar sun position (row 4), and weighting (row 5):

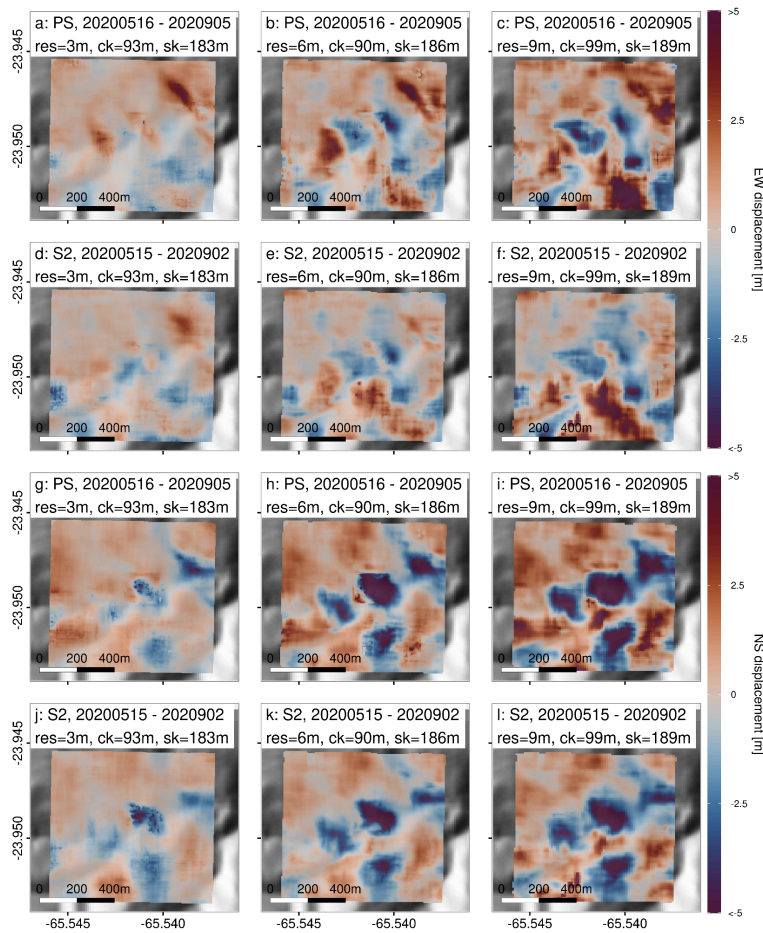


We observe that in the well-illuminated example (area I, left Figure above), the small acceleration is captured very well, and as the seasonal oscillation bias is low, little to no correction is required. In the topographically more complex area suffering from shadow effects (area II, right Figure above), the reconstruction is much more difficult. Here, the best-fit sine per pixel correction achieves the best results but could potentially smooth out a regular seasonal response by the landslide. For the Del Medio landslide, we would assume an intermediate case: the seasonal oscillations are higher than in test area I, but unlike test area II, with a similar aspect and less complex shadows.

We have added these additional analyses to the Supplementary Material (Figures S23 and S24).

This question also relates to the terrain features the correlation algorithm is sensitive to, depending on the optical sensor (resolution, noise), the parameters of the correlation algorithm and the type of land-cover. My feeling is that for a similar area of study, the high- and medium resolution images are sensitive to different features of the terrain. Your results on the different sensors could perhaps elucidate this question?

We have expanded our analysis of the bias effect at different resolutions to Sentinel-2 as well. This also allows us to compare downsampling a higher-resolution image (PlanetScope) and upsampling a lower-resolution scene (Sentinel-2). We used two scene pairs with very similar acquisition times (2020-05-15 and 2020-09-02 for Sentinel-2 and 2020-05-16 and 2020-09-05 for PlanetScope). The results are comparable and were added to Figure 5 in the revised manuscript:

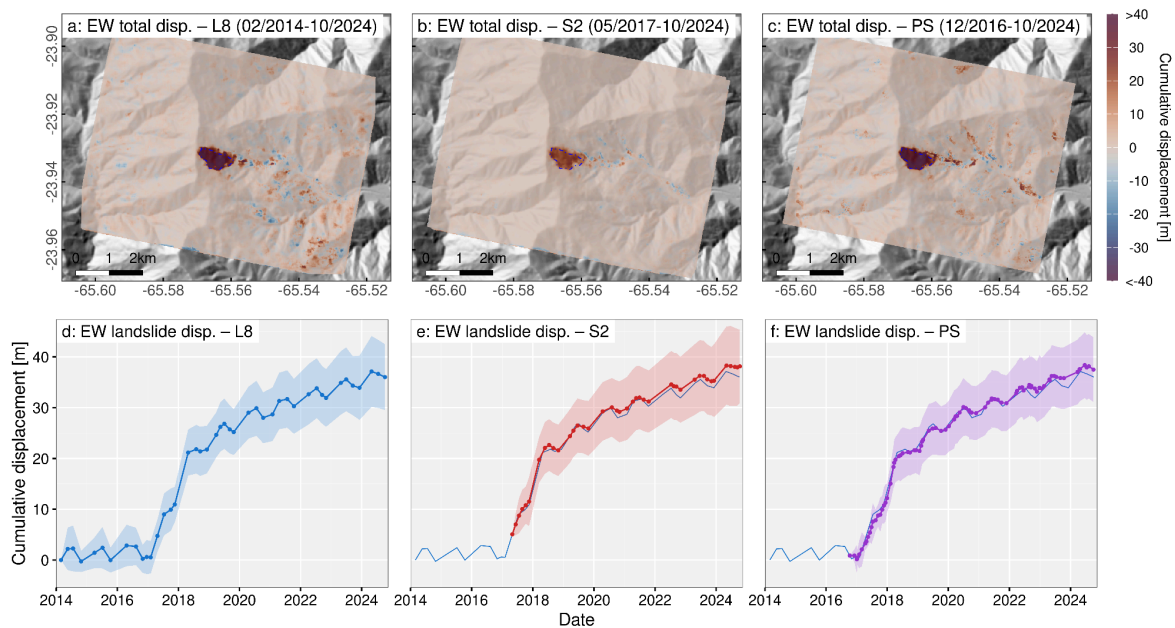


The results obtained with different sensors are really intriguing as S2 and PS present lower displacement values on the Del-Medio landslide than with L8 (Figure 2), despite their higher resolution. This is counter-intuitive (cf e.g. Bontemps et al., 2018 ; Cusicanqui et al., 2025), and questions me a lot on the ground of these results. Is that also an effect of the terrain-feature depending on the sensor type (or resolution)? That would be really interesting to apply your method to another area with validation data (or synthetic data). In the supplementary material, Figure S17 shows a result of synthetic tests, but it is not clear how it



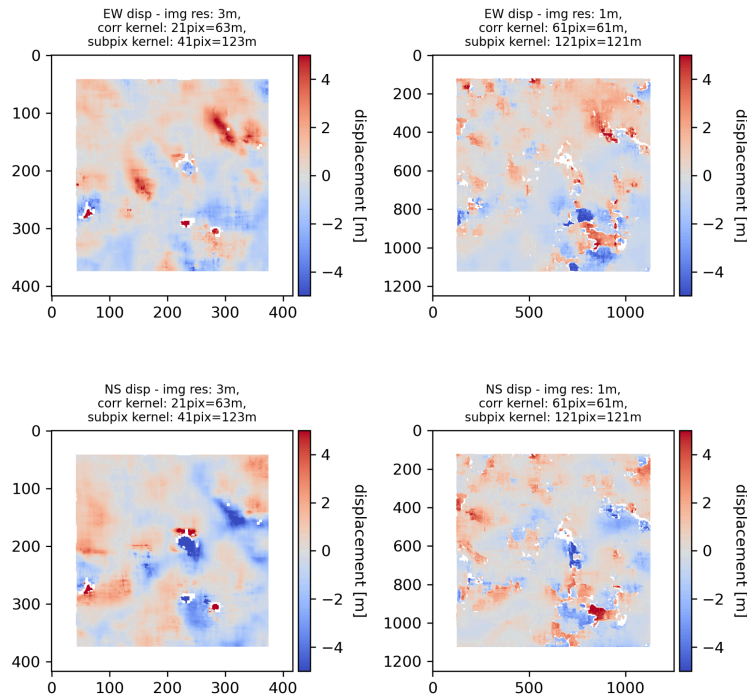
has been developed and what are the results for the different sensors. Maybe it could be worth focusing on that aspect?

It seems like the slightly lower velocities for Sentinel-2 and PlanetScope were linked to noisy pixels around the landslide border (i.e., correlation border effects). Following your suggestion (see comment below), we have applied an inward buffer of half the correlation kernel size to the landslide mask to exclude bias from border effects. The results between all sensors are now more comparable. Here is the updated Figure 2:



The results on the impact of the sensors presented in Figure 5 (and section 4.3) investigate the effect of the spatial resolution only, whereas the satellite images have not only their own pixel resolution but also their own pixel noise (radiometric noise). A pixel from a PlanetScope image is much noisier than a pixel of S2, so that modelling a S2 image by downsampling a PlanetScope image at 9m (~10m) resolution is certainly not valid. To me this paragraph does not allow you to elucidate the differences observed between the different sensors. I would rather try to compare correlation results obtained with different satellite images taken at similar days/hours of acquisitions. Nevertheless, I found this part interesting as it made me also wonder « does a PlanetScope upsampled at 1m lead to even lower seasonal oscillation ? ».

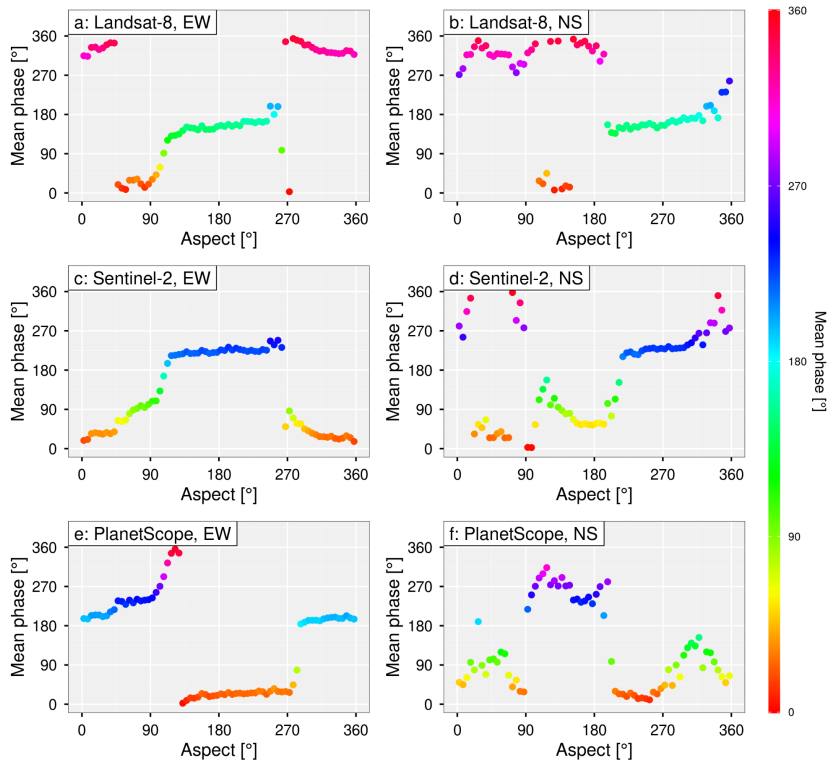
Running the process with 1 m resolution and still keeping the correlation kernel size constant in terms of area is computationally very heavy for even this small area, making it unfeasible to run, especially for larger areas. What we did, however, was to compare PlanetScope 3 m to 1 m resolution, but using a smaller correlation kernel of approximately 60 m. Again, the magnitude of bias from different illumination conditions is reduced, but due to the overall smaller size of the correlation kernel, the results appear less smooth and have more gaps:



We have expanded Figure 5 by adding a comparable analysis using Sentinel-2 data from a similar date and updated the section accordingly.

Some technical aspects must be clarified : (i) no phase delay on the modelling of the seasonal oscillation appears in the Equation 9, despite a map of phase delay is shown in Figure S8 and S10. How is the phase delay taken into account? This phase delay is, by the way, not analyzed at all in your results/discussion. I think that's a bit of a shame. (ii) the definition of « short-baseline » pairs in relation with inter-annual trend and the choice to not used these « short-baseline » pairs in the inversion strategy are not clear. (iii) how is done the upsampling of S2 and L8 images?

(i) The phase was included, but the term was missed in Equation 9 – thank you for spotting that. We have corrected the equation. We have also included a small discussion on the phase delay together with the following plot to the Supplementary Material, which shows the mean phase delay per aspect bin (5 degrees).



We observe a clear phase shift for south-west facing slopes in the EW component, which is again the direction of shadow extension, while for the NS component, the pattern is more diffuse. Interestingly, phase delay is similar for Landsat-8 and Sentinel-2, while it seems inverted for PlanetScope. This may be related to the higher variability in sun position, particularly sun azimuth in the PlanetScope scenes used.

(ii) We excluded short-temporal baseline pairs when we expect the landslide motion to be below the detection limit. This should not affect the seasonal bias component, since illumination biases are introduced anyway when cross-season pairs from different years are correlated.

(iii) The upsampling is done through cubic interpolation as mentioned in line 117 (initial manuscript).

I found that the section « 4.4 mitigation of seasonal bias » must appear in the methodological section rather than in the results. One key aspect of your study is a comparison of the mitigation strategies of the seasonal bias. This different strategies should first appear in the methods, before showing the results.

We have moved the description of the correction strategies to the methods section in the revised manuscript.



The introduction/title/abstract are too much focused on the Del-Medio landslide. whereas this single-case study is just an example of application of your methods to one specific case-study. To my point of view, the main brought of your work are elsewhere, on the comparison of the seasonal bias mitigation, including for a multi-sensor approach. I think the introduction should be partly rewritten, to decrease the parts related to the Del-Medio landslide (for instance the text from lines 42 to 48 could be placed in the « site study » section), and to bring more elements on the state of the art on the multi-sensor approach for ground-motion measurements.

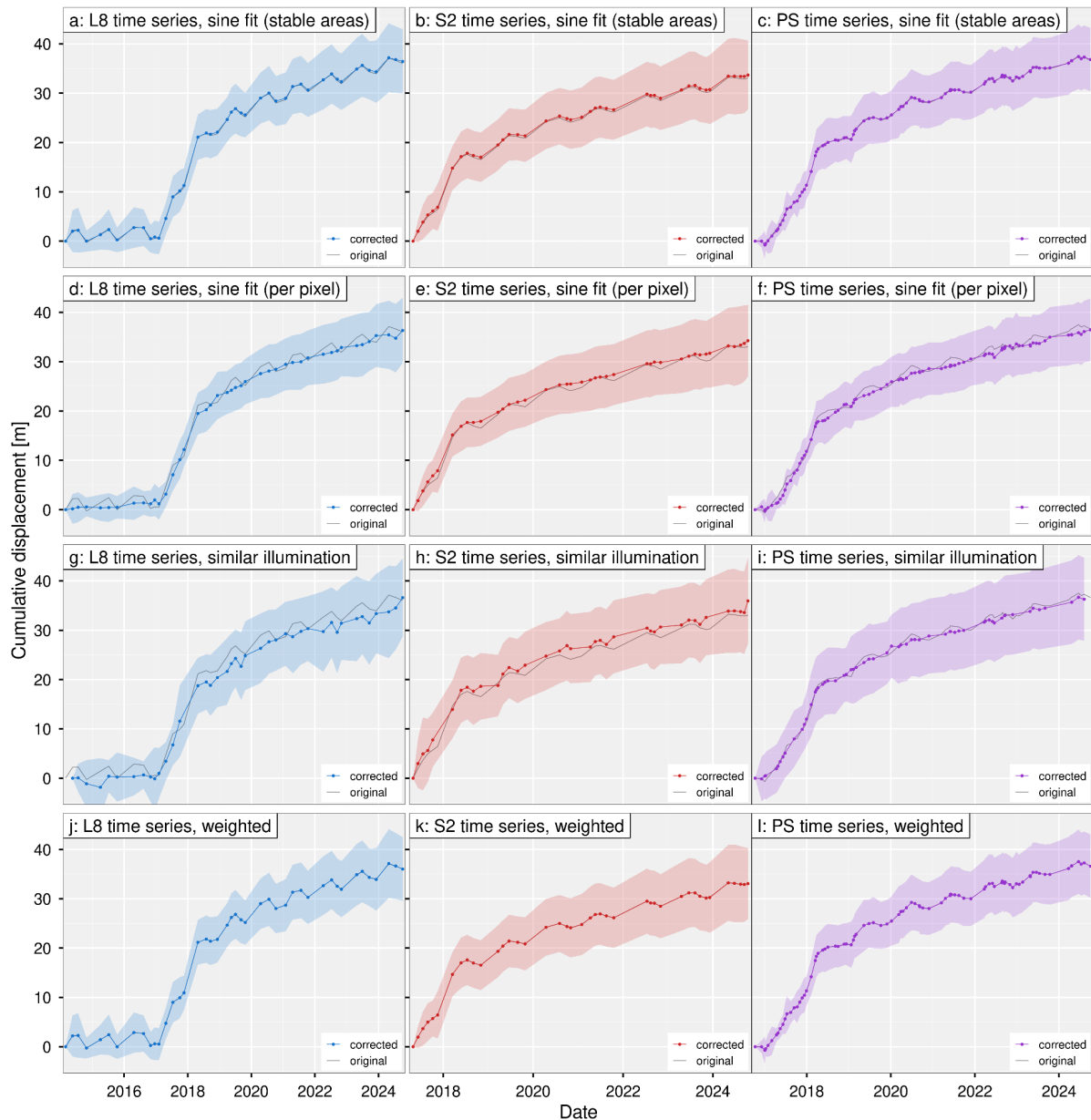
We have restructured the introduction to better separate the general problem of changing illumination for optical offset tracking applications from the case study of the Del Medio landslide. Nevertheless, we do believe the Del Medio landslide is a great example of the problem that the illumination changes introduce: We have a landslide situated in a seasonal environment, making true responses of landslide motion during the rainy season likely. At the same time, seasonal illumination changes and moving shadows from the steep surrounding terrain introduce a severe bias effect. Can these two signals be untangled, and if so, how? – This is the main motivation of this work.

The title/abstract should also reflect this aspect of the study, and provide less importance to the Del-Medio results given that no validation datasets are existing here.

See our response to the previous point. We did not change the title of our manuscript, as we believe it describes well the two aspects that are covered in our study: (a) challenges of and mitigation strategies for seasonal biases, and (b) the Del Medio landslide, which is analyzed in detail. We now include a synthetic analysis that takes advantage of the complex relief of the study area.

The results obtained on the Del-Medio landslide with and without seasonal bias mitigation using the L8 sensor (Figure 7) are not intuitive to me. Indeed removing the seasonal bias leads to a much lower L8 total displacement. This is not the case for S2 and PS. How would you explain it?

This was also related to noisy pixels around the landslide border (see also the other comment related to the border effects). Pixels impacted by border effects were now excluded through an inward buffered mask, which reduced standard deviations and made measurements across all sensors more comparable. Below is the updated Figure 7 with the more restrictive mask applied. We have also added the displacement time series from the two other correction approaches for completeness, although they very closely align with the uncorrected signal:



Paragraph 4.5 is not really clear to me:

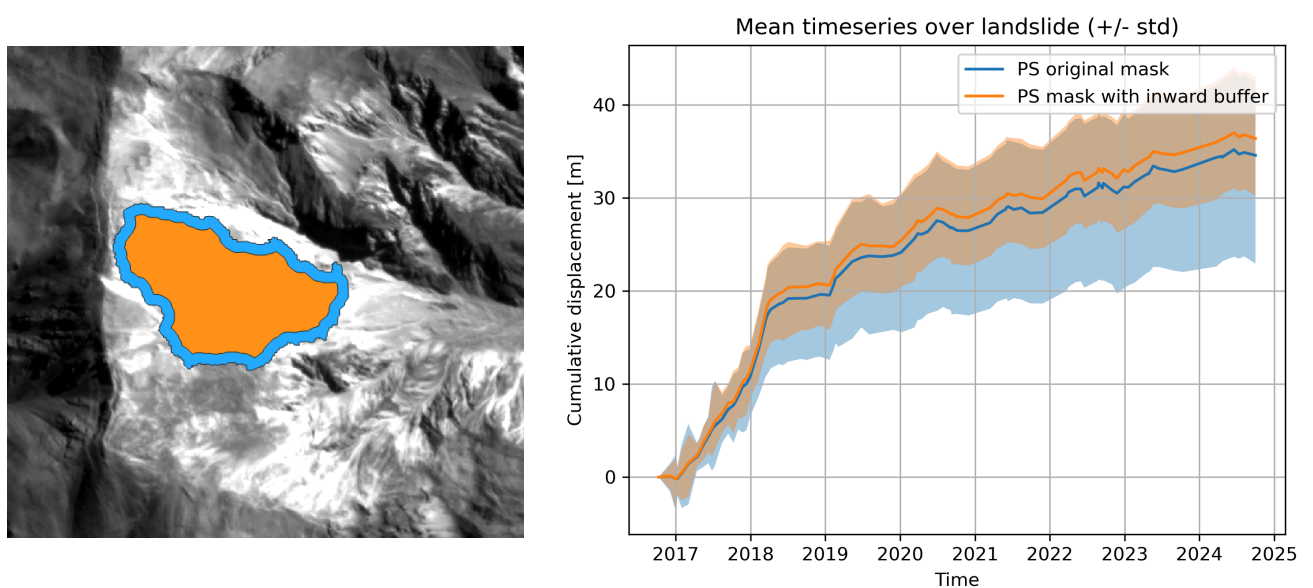
The « high » standard deviation observed in the Figure 7 is not only linked to difference of unit kinematics. If I understand well, the ribbons shown in the Figure 7 reflects the standard deviation of all the pixels. To me this pixel variations is linked both to the kinematics of the different units and to border effects due to size of the correlation windows (See for instance Bontemps et al., 2018).

The standard deviation of all pixels over the landslide is shown. Therefore, if some pixels move faster and others slower, the standard deviation becomes larger.

Also I am wondering the effect of averaging cumulative displacement time-series over areas not showing the same velocities. I don't think this is the best way to highlight the kinematics

of the landslide. I would rather pick up some selected pixels in the middle of the different units of the landslide (to limit the border effects very common with medium resolution satellites). This comment could also apply to the time-series of Figure 10.

The reviewer is correct in pointing out potential border effects where the correlation kernel covers both stable and moving ground. We estimate that border effects would concern pixels within a perimeter of half the correlation kernel size from the landslide outline, which is  $\sim 17$  pixels  $\times$  3 m resolution = 51 m. We created an inward buffer of the original mask to remove these pixels. The effect on the mean cumulative displacement time series is small; however, the standard deviations and also the lower displacement values are greatly reduced. The measurements between the different sensors are now more comparable. Here is the before and after comparison for the PlanetScope timeseries:



To showcase the spatial variability, we refer to the profile plot where the cumulative displacement along an EW transect is shown.

It is also not really clear what does the term « uniform acceleration » refer to?

With this, we mean that the displacement estimated over the entire landslide body was similar (uniform) for the 2017 event. Later, different parts of the landslide, e.g., areas close to the upper scarp, moved faster than the lower part of the landslide. We have made a small addition to the sentence to clarify this:

“The high standard deviations following the relative spatially uniform acceleration across the entire landslide surface in early 2017 [...]”

The word « surge » is used a lot in the glacier community but is not really common for the landslide community. It is sometimes used to describe specific events like landslide-induced-tsunamis. I don't think it is the right word to use here (and elsewhere in the text).

We have replaced all occurrences of “surge” with “acceleration”.

The accelerations from 2018 and 2021 are not really obvious to see in Figure 7, also given that Figure 7 averages pixels with different velocities (and maybe even different timing of accelerations of the different units). Maybe try to better highlight these accelerations in the mean time-series normalized per-pixel of individual time-series in the fastest parts of the landslide where the signal to noise ratio will be higher?

Yes, averaging over the entire landslide smooths the 2018 and 2021 responses. However, in Figures 8 and 9, the accelerations are better visible based on the profile line and a smaller kernel in the upper portion of the landslide. The objective of Figure 7 was primarily to show the result of the correction approaches.

The analysis of the link between the landslide accelerations and the rainstorms can be improved (after these 2018 and 2021 accelerations have been better highlighted. See my previous comment). First of all the brought of ERA5, due its low quality in this area, is very limited (as also pointed out by the author). I don't see the point of showing it. I would keep only the GPM dataset. However, I found interesting that extreme rainfall events could accelerate deep-seated slow-moving landslides. This is not the first case-study showing this. I don't know the complete bibliography, but I have in mind the Joshimath landslide (Sreejith et al., 2023). A better analysis of the litterature can be done and the mechanics behind this acceleration analyzed.

We have removed ERA5 total precipitation, but kept the ERA5 temperature because these could be relevant for freeze-thaw processes. We have also extended our literature review to include additional references.

The Figure 10 is really confusing to me: Is that a total displacement or the amplitude of the seasonal acceleration ? The caption and the legend says different things. Since I didn't understand this Figure, I can't really comment it as well as the section 5.2 that relies on this Figure. Please clarify and I could comment during a second round of review.

The legend label was wrong and was corrected – thank you for spotting this. What is shown is the difference in total estimated displacement between the three sensors for a common time period for both the corrected and uncorrected results. We also modified the Figure labels to better highlight this.

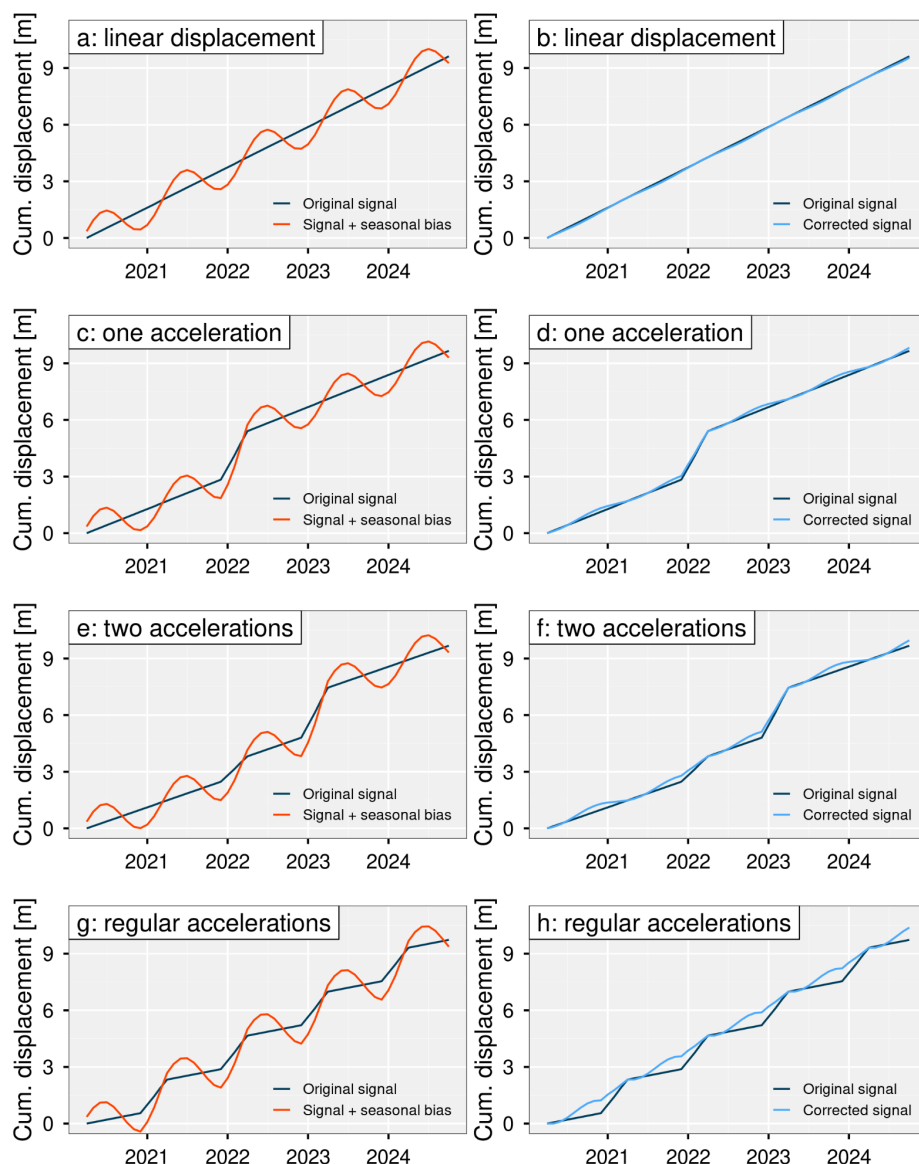
In Figure 4, it would be interesting to have a probability density function of the slopes (both slope gradient and aspects) and the sun-azimuth differences. Indeed for slopes or sun-azimuth difference appearing few times, the results are certainly too noisy, and the bins must be removed.

A Figure that shows how many measurements fall inside each bin is shown in the Supplementary Material, Figure S12.

If I understand well, for the specific case of the Del-Medio landslide, the seasonal oscillation has been removed per pixel (lines 270-278). As the authors say, this can be applied to the Del Medio landslide only as it seems to not show any seasonal movement (even if the absence of in-situ measurements this assumption is complex to validate). I think it must be even more clearly stated that this method should not be applied to any landslides.

The reviewer's understanding is correct – the fit was applied per pixel. This method is suitable if a landslide does not show a regular seasonal displacement component, but only for more occasional acceleration phases, as we interpret the displacement history of the Del Medio landslide. Having a time series of many years is also helpful in this regard. We have experimented more with different synthetic displacement scenarios to better constrain when this method can be applied and when it introduces a bias of its own: linear displacement and signals with irregular acceleration phases (i.e. not every year) can be reconstructed well, while a regular seasonal landslide response would be smoothed out (we replaced former Supplementary Figure S17 with this updated analysis):





For the Del Medio landslide, because the results of the sine fitting approach largely align with the reduction of seasonal biases by only using correlation pairs with a similar illumination, we assume that the approach is relatively safe to use. However, due to a lack of ground-based data, we cannot conclude this with absolute certainty. We have added this constraint more clearly in the revised manuscript and base our analysis of the landslide displacement history on the data obtained from similarly illuminated pairs only. This indeed revealed two more potential small acceleration phases that were not visible when subtracting the best-fit sign per pixel

Overall, every correction method has its drawbacks and alters the observed signal. E.g. with the sine fit over stable areas with similar exposition, we observe that the oscillations over the landslide were not significantly reduced – is keeping the illumination bias and not being able to correctly interpret the landslide’s seasonal behavior better than partially removing it but effectively mitigating bias? This is a tradeoff that has to be considered on a case-by-case basis, which we have stated even more clearly in the revised manuscript.

The choice of not using the short-baselines must be better discussed. Indeed it is certainly a good choice in the case of the Del Medio landslide where the land-cover doesn't seem to evolve much with time (maybe I am wrong?), but this might not be a good choice in other areas with a human presence and modification of the land-cover with time, that can lead to a decay of the coherence with time. This must be discussed.

Indeed, the land-cover changes in the Del Medio catchment are limited as there is barely any anthropogenic impact at that elevation, and vegetation cover is sparse. The only rapid changes occurring are observed in the alluvial river bed, which sometimes can complicate the matching in this part of the basin. We have added a sentence to consider this aspect in study areas with more dynamic landcover changes.

In Figure 8, we still observe seasonal oscillations, despite a per-pixel seasonal signal has been removed (if I understand well). This is not really intuitive to me why these oscillations still remain? Also the transect AA' seems purely EW oriented, and not exactly longitudinal. Wouldn't it be better to use a longitudinal transect of the landslide?

Yes, some oscillations are still present, but their amplitudes are greatly reduced in comparison to the original measurements. The sine fit will always have a regular amplitude, which will not perfectly fit every oscillation especially in the presence of noise.

Considering the additional analyses we did based on synthetic displacement signals and the potential bias introduced to the time series through the per-pixel sine fitting approach, we agree that it is probably safest to use the time series retrieved from the similarly illuminated image pairs for subsequent interpretations, even though the overall lower number of correlation pairs make in more noisy for the Sentinel-2 and Landsat-8 time series. We have replaced the Figure showing the displacement along a profile for the sine fitting approach with the one from the similar illuminated pairs (previously Figure S20 in the Supplement).

We chose an EW transect because we focused on EW displacements only (to not mix seasonal bias effects from EW and NS).

In Figure 10, the standard deviations of the difference of displacement fields on stable areas between L8-S2, L8-PS and S2-PS are calculated. This can be exploited to estimate separately the standard deviation of the displacement fields of L8, PS, and S2. Given the values shown in Figure 10, it will show  $\sigma_{S2}$  and  $\sigma_{L8}$  higher than  $\sigma_{PS}$ , with  $\sigma_{S2}$  and  $\sigma_{L8}$  relatively similar. In this context I am wondering why S2 and L8 do not show similar time-series? Also, adding these standard deviations to the individual time-series (for instance on Figure 9a) would provide a more convincing argument to the analysis of the accelerations period.

We believe the reviewer refers to the possibility and usefulness of investigating the standard deviation of differences between L8, S2, and PS at every time step, not the total cumulative

displacement estimated (as currently shown in Figure 10)? However, this would require the time series to be interpolated to a common sampling interval because the acquisitions are not from the same date across all sensors, making it less intuitive and also less of a fair comparison. After all, if the data are linearly interpolated to a regular monthly sampling interval, periods with low actual data coverage (rainy season) would be overrepresented. We therefore kept the initial layout.

Some Figures could be partly improved. In Figure 3, the dashed black contour of the landslide is not well visible. Maybe choose another color or symbol? In Figure 9b the dashed black line indicating the accelerations mask the precipitations. In Figure 9a, indicate more clearly that the major debris flow is not occurring on the landslide but nearby.

We have implemented the suggested changes. For Figure 9, considering the potential bias the per-pixel sine correction method introduced, we changed the time series used for interpretation of the landslide behavior to the results obtained from the images with similar illumination, also with regards to the data shown in the profile plot. The strong reduction of displacement fields also reduces the number of available dates which makes it more difficult to constrain the precise onset of an acceleration phase. We have changed from indicating the increase in displacement rates through a single line to highlighting a wider period between the two (or three if onset is not very clear) closest available satellite images.

Using similarly illuminated pairs only, the PlanetScope time series at the upper landslide scarp appears less noisy than the one of the per-pixel sine fit (initial manuscript). Based on the new plot, one may even argue for two additional small acceleration phases between September to November 2019 and January to March 2020. These are, however, only present in the PlanetScope time series (temporal sampling of Landsat-8 and Sentinel-2 is too coarse with the sun elevation and azimuth thresholds applied). We have therefore added these phases in light gray to the plot and adapted the manuscript accordingly. Here is the updated Figure:

