Preliminary response to reviewer 3

We thank the reviewer for taking the time to review our paper. We will prepare a full response to all comments made on the manuscript later on in the review process, but here we provide a preliminary response to some of the major points raised in this review (reviewer comments in bold)

Generally, I suggest to report the results more as a potential contribution towards using Sentinel-1 data for estimating time-windows of event occurrence. Some parts of the manuscript read as if a well-working method is presented that works generically for identifying timings of landslides. However, this is still very much work in progress. For instance, I don't think that *"This will allow multi-temporal landslide inventories to be generated for long rainfall events such as the Indian summer monsoon"* in a comprehensive manner.

We will revise the manuscript to make it clearer throughout that our methods cannot establish timings for all the landslides in an inventory (and in fact, will only provide timings for ~20%).
The quote here is taken from the abstract. In response to this and to similar comments made by reviewer 2, we will change the end of the abstract.
Previous text: *"our methods allow 20% of landslides to be timed with an accuracy of 80%. This will allow multi-temporal landslide inventories to be generated for long rainfall events such as the Indian summer monsoon, which triggers large numbers of landslides every year and has until now been limited to annual-scale analysis."*
New text: *" our methods allow 20% of landslides to be timed with an accuracy of 80%. Application of our methods could provide an insight on landslide timings throughout events such as the Indian summer monsoon, which triggers large numbers of landslides every year and has until now been limited to annual-scale analysis*" .
This removes the words "multi-temporal landslide inventories", which were misleading since we cannot provide a complete inventory where all the landslides have timings assigned.
We will make also make sure this is clear in the discussions / conclusions and throughout the manuscript.

We will also make sure it is clear throughout that we are not assigning specific dates, but instead time windows of (in most cases) 12 days to each landslide.

There will definitely be a biases in terms of identified slides,
The biases towards which slides can be assigned a timing is related to section 3.3 "Factors effecting the performance of each method." For example, it is clear that larger landslides are more likely to be assigned a timing than smaller landslides using our methods (Fig. 5 d-f).
In the manuscript, we considered how this effected where our methods could be applied (for example, they will not work well for inventories of small landslides or in arid environments), but we did not consider how well the timings of the 20% of landslides we assign a timing to using SAR methods will represent the full inventory. In fact, this 20% will be biased to contain a higher proportion of large landslides and landslides in more heavily vegetated areas than the original inventory. We will add this point to the revised version manuscript

a vast majority of sildes will be missed or - worse - labelled incorrectly,
Our methods should not incorrectly label a large percentage of slides. We expect that if we apply our methods to an inventory of rainfall-triggered landslides, we will obtain an inventory in which ~80% have no timing information, ~16% are correctly timed and ~4% are incorrectly timed. We will be careful to make this clearer in the abstract, results and conclusions sections of the revised manuscript.

and things might look dire when thinking beyond the scope of this study, e.g. if no polygons are availabe.

If polygons are not available for an event, it will not be possible to apply our methods in their current form – they are not designed to be applied to events for which we do not have a pre-existing landslide inventory. We will ensure this is clear in the revised version of the manuscript.

It took me a while till I figured out the meaning of the terminology you used for the orbit IDs (e.g. "H083A"). Please specify more clearly that this is a combination of study area, orbit number and orbit direction.

Yes, this was also raised by reviewer 2, who suggested a change of terminology from e.g. "H083A" to "Hiroshima_asc" to describe the tracks. We will make this change, which should also resolve this comment.

"We tested both of these polarisations, but found VV to perform better than VH so present only the results for VV." This is an interesting finding. How was this evaluated?

Yes it is interesting since other studies (e.g. Handwerger et al. 2022) use VH and find it to be the most successful option. We attach a version of table 2 in the original manuscript which contains the same results from VH in Zimbabwe and Hiroshima (VH data are not available for the Nepal case study, which occurred soon following the launch of Sentinel-1 before dual-pol data began to be consistently acquired – this would be a further disadvantage to using VH).

	Hiroshima		Zimbabwe		Trishuli		Buri Gandaki	Bhote Kosi		Hiroshima (VH)		Zimbabwe (VH)	
Track	H090D	H083A	Z079D	Z072A	Tr019D	Tr085A	BG019D	BK121D		H090D	H083A	Z079D	Z072A
Total landslides	543		383		650		922	1554		543		383	
non-masked	543	540	383	383	485	474	592	894		543	540	383	383
ls-b inc	137 (26%)	92 (36%)	90 (50%)	66 (41%)	106 (38%)	107 (35%)	152 (36%)	313 (36%)		94 (12%)	97 (29%)	40 (20%)	26 (12%)
ls-b dec	126 (38%)	205 (57%)	182 (23%)	155 (34%)	156 (37%)	143 (22%)	113 (27%)	310 (32%)		182 (25%)	251 (63%)	262 (28%)	266 (36%)
pix var	160 (45%)	181 (59%)	134 (50%)	83 (47%)	141 (42%)	125 (44%)	141 (30%)	261 (43%)		155 (32%)	222 (56%)	152 (16%)	110 (27%)
shadow	79 (49%)	122 (75%)	43 (55%)	58 (72%)	45 (80%)	50 (80%)	17 (88%)	52 (87%)		144 (43%)	227 (74%)	125 (39%)	140 (47%)
combined >2m	51 (75%)	99 (88%)	48 (65%)	39 (85%)	45 (76%)	33 (88%)	33 (64%)	86 (63%)		71 (54%)	166 (80%)	88 (33%)	87 (60%)
combined >3m	10 (80%)	22 (100%)	5 (80%)	2 (50%)	8 (100%)	3 (100%)	1 (100%)	5 (100%)		9 (89%)	38 (97%)	17 (35%)	9 (89%)
asc & desc	122 (80%)		81 (73%)		70 (81%)		-	-		196 (69%)		142 (44%)	
asc & desc 2m,1t	79 (73%)		64 (72%)		52 (77%)		-	-		139 (65%)		102 (23%)	
asc & desc 3m	43 (91%)		17 (76%)		18 (94%)		-	-		57 (81%)		40 (75%)	
baseline (1/n)	7%	17%	10%	7%	8%	8%	8%	14%		7%	17%	10%	7%

As you can see, we have a generally lower accuracy for VH data, especially in Zimbabwe. Since this may be interesting for future studies, we can include it as a supplement.

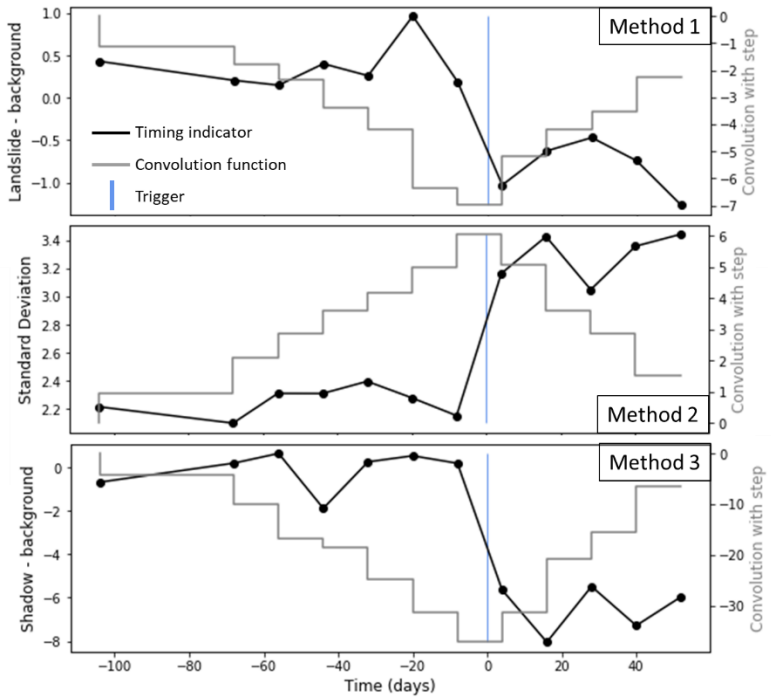
The copernicus DEM would have been a more recent DEM version, also available globally at a resolution of 30 m.

We chose to use the SRTM DEM here since it is already available as a dataset in Google Earth Engine, making it easier to integrate into the slope correction module for future users.

l. 161/Figure 2: "A step change in the difference between the median landslide amplitude and the median background amplitude is then used as an indicator of landslide timing." It might be beneficial to plot this difference?

Yes, this is a good point. In response to this and to similar comments from reviewer 2, we have prepared 3 panels (below) which will be incorporated into Figure 2. This makes the step change clearer (Note we have also changed to a different landslide polygon due to changes requested by reviewer 2 – these time series are thus not comparable to those in Figure 2 of the original manuscript).

These panels also show as a grey line the convolution between the method and a step function to make it clearer how a step change results in a peak or trough in the convolution function.



Overall, appropriate performance metrics and their interpretation is of key importance. In fact, when thinking about the implications of the method presented here, this is crucial. If no validation data are available (e.g. when this method is applied to a new data set), a vast majority of identified dates (more precise: time windows) will be incorrect. This needs to be discussed.
The problem is not that a vast majority of landslides will be assigned an incorrect time window, but instead that the majority of landslides will not be assigned any time window at all. The subpart of the inventory that has timings assigned will be biased towards larger landslides and those in more vegetated areas. We will discuss this in the revised version of the manuscript. (See response to earlier comments.)

Of those landslides that are assigned a time interval, we expect that 80% of the time, this interval should be correct. Therefore if, in a new dataset, we observe a spatio-temporal cluster of landslides, we can assume the timing of this cluster is correct, since it is very unlikely that all the timed landslides in this cluster would be assigned the same incorrect date.

Publishing the code (e.g. on GitLab/GitHub) would be welcome for the final manuscript, but also of interest from a reviewer's perspective. If there are concerns with respect to sharing code before the publication is accepted, there are surely opportunities for embargos.

The GEE and python codes will be shared as a supplement to the next version of the manuscript.