

Nat. Hazards Earth Syst. Sci. Discuss., referee comment RC1
<https://doi.org/10.5194/nhess-2021-18-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on nhess-2021-18

Jan Henrik Blöthe (Referee)

Referee comment on "Timely prediction potential of landslide early warning systems with multispectral remote sensing: a conceptual approach tested in the Sattelkar, Austria" by Doris Hermle et al., Nat. Hazards Earth Syst. Sci. Discuss., <https://doi.org/10.5194/nhess-2021-18-RC1>, 2021

Review of Hermle et al.: "Challenging the timely prediction of landside early warning systems with multispectral remote sensing: a novel conceptual approach tested in the Sattelkar, Austria".

Manuscript Number: nhess 2021-18

In their manuscript, Hermle et al. propose and test a conceptual approach for landslide early warning that seeks to make use of optical imagery with high temporal resolution to ensure a timely warning to be issued. They term the time between the issued warning and the failure "lead time" and argue that high revisit times of novel satellite systems and/or the immediate production of orthoimages from UAS surveys enables the researcher to detect a significant acceleration of mass movements within a time window of ~20-30 hours (under ideal conditions), thus making the "lead time" sufficiently large. Using image cross correlation to identify significant displacement between sequential imagery, they show that for the Sattelkar case study and the two historic examples of Vajont and Preonzo, their workflow could allow for a timely warning based on optical image assessment.

Overall, the manuscript is very well written and, in my view, represents a valuable contribution to landslide early warning research that merits publication in nhess. However, before acceptable for publication, the authors need to address number of concerns, mainly regarding the image correlation results and error assessment, which I will outline in detail below.

General comments:

A) Description of digital image correlation method and error assessment

In my view, digital image correlation is not a trivial method and deserves a more detailed description in section 4.3. Especially because the conceptual approach presented here grounds on the detection of significant movement (or even acceleration) from optical imagery, the authors should elaborate the exact processing steps and include a detailed accuracy assessment. This can easily be achieved by:

- The quantification of a level of detection between images, i.e. the residual mismatch of stable surfaces outside the landslide between consecutive images after image correlation, beyond which significant displacement can be detected with a given confidence.
- Excluding spurious matching results (displacement vectors) on the basis of a correlation threshold.

In the specific comments below, I detailed the exact locations where I think this would be important to address – I apologize for any repetitions.

B) Result of image correlation

As stated above, digital image correlation and the extraction of displacement from correlated imagery is not a trivial task and many pitfalls can lead to spurious results (the authors term these decorrelated). I will outline my doubts regarding the validity of the obtained displacement values referring to Fig. 5, but have given many detailed comments on the respective text positions in the specific comments below. In large areas, the image correlation returns areas that are “decorrelated”, such as the western part of the landslide in (a) and (b), but also positions in (e) and (f) are affected by this. In my experience, such a pattern indicates that matching between images did not work, which should be visible by adjacent vectors having very different magnitudes and directions. Furthermore, the patchy nature of displacement values in the western part of (c) is very surprising. Here very high total displacement of ~18 m is located in the vicinity of displacement on the order of 4-8 m. From an image matching procedure, I would expect a rather smooth picture here, such as in (d). But also from a geomorphic perspective, I am unsure how this pattern could be explained by a natural process. Finally, the results obtained from the downsampled UAS DEMs predominantly show high rates (16-18 m) that are interrupted by areas of no movement or very slow movement. My impression would be that these results are least reliable, because a) they show a completely different picture as (a) and (b), while being computed with the same data (just a different resolution), b) the displacement values are nearly the same for two very different time intervals (e = 376 days, f = 42 days), c) they are not matching the values obtained from manually tracking boulders (again, based on the same data), and d) I am unsure if such a pattern can be produced by a natural process.

Having outlined my reservations regarding the image correlation results, let me suggest a couple of strategies to improve the results:

- Use a hillshade not a DEM for tracking (not clear if this was done)
- Resample the DEM to a slightly coarser resolution (0.5 m?)
- Try a different software for image correlation, there are many and all have their advantages and disadvantages
- Have a detailed look into the correlation coefficients and the bearings of the displacement vectors and exclude spurious results.

Specific comments:

- L22/23: While this is certainly true, the authors should elaborate in the introduction that events instantaneously triggered by earthquakes or heavy precipitation are beyond what their proposed framework can deliver an early warning for. The necessity of gathering and evaluating data prior to issuing a warning limits the analysis to mass movements that indeed show a pre-failure acceleration on the order of days.
- L25/26: Is this really just attributable to the warming of the climate?
- L47/50: I would think that also the rate of landslide movement defines whether or not it can be detected by optical imagery.
- L79/80: This is the maximum revisit time at the equator, right? For the study area shown here, revisit time should be shorter.
- L121: What do you mean by "natural developments" and how are these conditioned or different from natural processes?
- Figure 1: While I like the idea behind this conceptual figure, I would recommend the authors add a time axis and limit the area of "significant acceleration" to a vertical line that coincides with $t = 0$. In the present form, the conceptual figure contradicts statements in the text, such as "The forecasting window is started [...] following significant acceleration [...]" (L126), or "Simultaneously with the forecasting window, time to warning (t_{warning}) starts (grey outline)" (L128/129).
- L133/134: This also does not match what Fig. 1 is showing
- L139: This also does not match what Fig. 1 is showing. In Fig. 1, $t_{\text{lead}} < t_{\text{react}}$.
- L215/127: In theory yes, but as you show later (Tab. 2), the effective revisit time of optical imagery might in fact be very similar.
- L242/248: It might be worth mentioning here that on average, only 11% of the images were usable, significantly reducing the theoretical revisit time, as you also outline in the discussion.
- L267/269: Please elaborate how you filtered for "errors of location, shift and spectral colour problems" (are the latter spectral differences between images?).
- L281/285: Please specify the accuracy of dGPS coordinates as measured for the GCPs and also include an accuracy information for the DEMs and their derivatives that were produced from UAS surveys.
- L285/286: Please elaborate how image co-registration was achieved and state here the residual mismatch between co-registered images.

- L288/289: Usually matching between consecutive images is not achieved by matching “common pixels”, but by maximizing the correlation between pixel-value distributions of patches of pixels (i.e. your windows of different sizes in Tab. 6).
- L304/305: What is the uncertainty of these east-west and north-south displacement estimates? Did you check whether the bearing of the displacement matches the general slope of the Sattelkar?
- L307/308 and L440/442: This seems a bit arbitrary. How did you determine a cutoff-value of 4m displacement? How did you distinguish outliers from non-outliers? What is the confidence of your estimates?
- L308/309: This contradicts the descriptions of Fig. 5a, where you point out that “ambiguous, small-scale patterns with highly variable displacement rates” (L332/333) dominate the western part of the mass movement.
- L311/312: I am not convinced that manually tracking boulders in the same images that were used for image correlation can verify the results of this correlation. You can use these data to check if manual and automated tracking give consistent results. Comparing manually tracked boulders from UAS imagery could however be used to compare against the displacement estimates from satellite imagery.
- L320: As you present total displacement for different time intervals here, not rates in distance per unit time, I would suggest changing the title here. Same is true for L326, L346 and L361.
- L335/336 and L366: Did you check the direction of displacement for the areas of small-scale patterns of ambiguous signals? I would suspect that these are very heterogenous here as well. It would also be worth looking into the quality information (correlation coefficients) for these regions.
- L397: For a comparison (and also for a better readability) you could convert your total displacement to average rates of m yr^{-1} or cm d^{-1} .
- L398/399 and L402/404: Given the large differences in total displacement between sensors and resolutions used for image cross-correlation, I do not think that you can make this claim. Please use an appropriate measure to quantify the agreement between manual boulder tracking and the three different approaches used for digital image correlation.
- L419/422: This might be the case, though you tested larger patch sizes (Tab. 6) that should have given you consistent results for this region then.
- L433/434: This should be backed by a statistical measure. From a close look to Fig. 5, I rather get the impression that the only patches you can make this statement for is location a in Fig. 5 (b) and (d) and location c in Fig. 5 (a) and (c), but to a lesser extent.
- L445/447: The size of the snow patches does not play an important role. The presence of snow in one image hampers correlation between images and leads to false patch-matching results.
- L457/462: To be frank, I do not see much similarity between Fig. 5 (c) and (e) nor (d) and (f). I would be very cautious in interpreting these results as is. This is especially true for the resampled UAS results.
- L463/464: As the GCPs for referencing the UAS data are probably located close to the landslide, it is not surprising, but neither disturbing, that false displacement clusters appear outside the area of interest.
- L468/470: Again, I would not trust the displacement estimates of the resampled UAS data. While it is true that your manual boulder tracking identified 2 boulders with displacement of 10 or more meters, the remaining 34 boulders show something different.
- L471/476: While it might be true that the results obtained from image correlation of resampled 3m UAS data are better (internally) correlated and show a more homogeneous deformation pattern, this does not mean that the result is correct. As I outlined above, I have serious doubts regarding the interpretability of this data, as there is no agreement with the manually tracked boulder velocities (except 2 boulders). Also, from a geomorphic perspective, I am not sure how you would explain a velocity

pattern where high velocities dominate throughout the entire landslide, but are speckled with lower to zero movement within (Fig. 5 e and f).

- L485/488: Did you evaluate the proportion of false-positive displacements to true-positive displacements and if so, how did you do this and can you please include this data? Based on the image correlation results shown here, you can make this statement, but I would be cautious to make a general claim on the usability of the data.
- L552/554 / Table 7 / Figure 9: I do like the idea behind this, where the authors show that their proposed workflow would enable a timely warning in the case of historic landslides. However, in the case of Vajont, I think you should include a critical factor. While it is theoretically true that a "forecasting window" would allow for your workflow to be completed well before the failure, the slow deformation of Vajont (35 mm d^{-1}) in the 30 days will be well below the level of detection of your image correlation analysis, if you collect an image directly after the onset of "significant acceleration". In order to be detectable, movement must have accumulated a critical distance before data collection of your workflow can set in (30 days = 1.05 m total displacement) – a factor that in my view would be important to include here. '

Technical corrections:

- L1: Landslide
- L103/105: Check grammar
- L185: Is this really the source the authors need to cite for the location map?
- L229: beginning of April
- Table 3: Here you use a different date format than in the text
- L257: UgCS-Software?
- Table 4: Unit for GSD missing
- L273/274: Add this information to Table 5 and delete here
- L299/300: I guess this is only relevant if you explicitly mention the image-processing times.
- L398: can be compared
- L409/410: resulting from significant morphological changes?
- L443: bracket missing?
- L 460: check figure reference