Minor revision of *Machine learning based nowcasting of the Vögelsberg deep-seated landslide: why predicting slow deformation is not so easy*. [egusphere-2022-950]

We were pleased to read the reviewer's appreciation on the improvements made to the manuscript. Here we would like to summarize the changes made based on the recommendations in the second review. Our improvements in the manuscript invoked by the comments, in brown, by the reviewer are detailed below. Independent of the reviewer's comments, the following textual changes were made to the manuscript:

- The colors of the lines indicating the deformation (ATS) in Figure 3 have been swapped to be consistent with the other figures in the manuscript.
- The ordering of the appendices, author contributions, competing interest statement, acknowledgments and references, as indicated by the editorial office.

- **Environmental conditions**

> **2.** I'm a bit concerned by the environment predictor data you use in you model. The spatial resolution of the satellite data is often much coarser than the size of the landslide. I have a hard time seeing how these coarse resolution datasets could provide any meaningful information at the local scale you're looking at.

We agree that the spatial-temporal resolution of the satellite observations is not optimal. However, these products are operational globally, and an important foundation for an operational nowcasting system. This is now explicitly mentioned in l. 59.

> **4.** Other studies (e.g., Thomas et al., 2019; Yatheendradas et al., 2019) have assessed the utility of satellite-based weather data for slope stability and found mixed results. I believe its possible that many of your issues could be attributed the poor representation of satellite data for your study site. I think this merits discussion or justification for why you think this is not an issue.

Satellite precipitation products are notoriously bad at measuring peak precipitation. The work by Thomas et al. (2019) and Yatheendradas et al. (2019), however, is on shallow landslides, that are more sensitive to this type of precipitation. We have added references to their work and a brief hint on the differences on l. 28.

> Additional point building on points 2 and 4 above. Since you have local precipitation measurements and a detailed hydrological model from Pfeiffer et al., 2021, I think it would be worth while to compare the satellite data to these more detailed data to see if that could be a major factor in poor model performance. As you say in the paper, "garbage in, garbage out".

Although this would be an interesting experiment we considered it to be out of the scope of this study. We now hint on this possibility and have made it explicit that we consider this out of scope of our study on l. 492. Our satellite focus is highlighted in all sections of the paper: in the abstract (l. 6, 12), introduction (l. 55), model variable selection (e.g., l. 189), and discussion (e.g., l. 502), and conclusion (l. 581). We changed l. 185 to further emphasize our desire for a satellite focus even though local deformation data is used in this study.

- **Previous work (Table C.1)**

> **1.** You provide a table of several research articles that have produced now-cast models. In section 2, you also say that at least some of these articles were for deep-seated landslides. I think it would be valuable for you to explicitly highlight what was different between this study and the ones that seemed to have success with now casting. Do you know exactly why they had success and you didn't? You discuss the reservoir being a factor for some of these studies, which makes sense. However, did any other studies try to nowcast deep-seated landslides with success?
>
> Additional comment on this. In your response letter, you say that the manuscripts you include in table C1 that don't have a strong driver only try and estimate smaller deviations from the trend. Does this suggest that they have significantly more data to develop and validate their models? Could this not be a major difference in your approach worth addressing explicitly? I'm not an expert on neural networks, but my impression is that they require much more data than alternative machine learning approaches. Using only 1500 data points to estimate ~68 parameters seems very problematic. I think you were getting at this in section 6.2.1, but a more explicit discussion is warranted.

To emphasize the dependency of many studies on a reservoir level, the references to previous work in Table C.1 has been split into two sections, with and without the reservoir level as input variable to the prediction. For readability, no 'track changes' version of the table and caption is provided.

Furthermore, we have expanded the text on this in l. 71 – 81. We now provide examples as well expand on our standpoint regarding trend splitting.

- **Minor comments and textual corrections**

> 23: Change 'global' to 'globally'
>
> 74: The last comma on this line needs to be deleted.
>
> 129: don't delete 'in'.

These textual corrections have been applied.

> 154: Their figure 3, right? Please clarify.

Indeed. This has been clarified.

> 158: Could you briefly mention what rainfall data Pfeiffer et al. Used?

This has been clarified as 'in-situ observations'. For further details on the full collection of local observations available the readers are referred to cited work.

> 199: Why 32 days? Move your explanation below up here.

There is no mention of 32 days in l. 199, this comment might reflect on l. 190 instead? The minimum filter length to remove all negative deformation rates induced by the noise was 32 days. This is explained before, in l. 162. The 32 day period was repeated for convenience here.

> Table 1: Consider defining API before this table or in the table.

Indeed, the API is introduced directly after rather than before the table. However, we prefer to provide an overview of the variables first. The API (V10) and seasonal noise (V11) require further explanation that would distract from the bigger picture of hydrometeorological variables.

> 224: What expert?

This has been made explicit as ‘experienced landslide hydrologist’.

> Figure 12: I would explicitly say the pink line is the mean deformation rate.

Adapted. Furthermore, the caption has been changed to clarify that this line indicates the mean of 1, 2 or 3 years of deformation.