

# **Interactive comment on “Deciphering seasonal effects of triggering and preparatory precipitation for improved shallow landslide prediction using generalized additive mixed models” by Stefan Steger et al.**

**M. Jurchescu (Referee)**

marta\_jurchescu@yahoo.com

## **“General comments”**

The submitted paper presents an innovative approach to combine the effects of precipitation on shallow landslides at four temporal scales: triggering precipitation relevant at the short-term, preparatory precipitation acting at the medium-term, cyclic seasonal conditions and across-year variability (the latter being relevant to account for biases in reporting landslide data).

The analysis is based on using the generalized additive mixed models (GAMMs) to account for the interactions among the precipitation variables and successfully separate between precipitation conditions leading to failures and those not inducing landslide movements.

The produced outputs are landslide occurrence probabilities associated to the interplay between short-term triggering precipitation and antecedent preparatory precipitation in dependence on the seasonal period of the year.

The seasonal effect is interpreted as being mostly due to temporal changes in vegetation and secondarily to temperature and snowmelt.

The manuscript is well structured and written and illustrations and tables are all necessary. In my opinion, the paper is worth publishing with only some very minor/technical revisions.

## **“Specific comments”**

A major strength and innovation of the study consists in the explicit consideration, the emphasis put on and the modeling of the seasonal imprints on the combined action of preparatory precipitation and short-term triggering precipitation in inducing slope failures at a daily scale.

The research is based on a data-driven modeling, applied following a thorough design of the study approach, and rendered highly transparently.

The robust development of the method, e.g. by eliminating the effects of year and locations, the careful and original design for sampling temporal landslide presence and absence observations at the mere landslide locations, and the data filtering criteria are well acknowledged.

The practical applicability of findings derives from the communication of predicted probability values also in terms of model performance metrics, such as *True Positive Rates* (TPR) or *True Negative Rates* (TNR), as well as from the welcome modes for visualizing (e.g. 2D and 3D plots, animations) the resulted thresholds varying throughout the year. The value of the research also consists in the spatial

and temporal cross-validation of the results, which ensure a high robustness and adaptability of the model to other regions and temporal periods.

The paper is also to be appreciated for the R-scripts and codes made available as well as the open data.

### “Technical suggestions”

- Section 3.1, page 6, line 160: if I understand correctly that the South Tyrolean data is part of the IFFI database, then I would suggest replacing “version” with “subset”;
- Section 3.1, page 7, line 165: please delete the plural “s” from the word “landslides”, since it should be the singular form;
- Section 3.1, page 7, lines 168-169: it is not clear what do you mean with “(incl. pre-2000 events)” in: “From January 2000 to the end of 2020, a total of 11420 points related to different movement types (incl. pre-2000 events) were registered for South Tyrol”;
- Section 3.2, page 7, line 188: please replace “times” by “with”: “multiplying the gridded daily anomalies with the gridded...”;
- Section 3.3, page 8, lines 206-207: please adjust for the repeating word “based” and insert commas e.g.: “For the cross-validation (CV), based on a leave-one-factor-out data partitioning (Section 4.5), ~~was based on~~ several spatial environmental variables were used (Fig. 2).”;
- Section 3.3, page 8, line 207: I would suggest replacing “this data” with “the data”: “The data was obtained from the open Geodata platform of South Tyrol (Geokatalog, 2021).”;
- Section 3.3, page 9, caption of Fig. 2: I would propose not using abbreviations in general in figure captions, thus I would propose replacing “CV” with “cross-validation”;
- Section 4.3, page 12, line 313: please replace “This study used a GAMM to discriminate precipitation...” with “In this study, a GAMM was used to discriminate precipitation..”
- Section 4.3, page 13, lines 343-344: please replace “between” with “among” in: “In detail, this *YEAR* variable systematically captures data variability among the single years”;
- Section 4.3, page 14, caption of Table 1: I would suggest adding in the caption information on the software used, e.g.: “Model setup and variables introduced into the binomial GAMM by means of/through the R software”; also, in the last row of the table, it would be clearer to specify “the following R command”;
- Section 5.1, page 16: lines 413-424 present the resulted number of records in the landslide samples as well as the impact of filtering out “dry” observations; however, for a better understanding and avoidance of repetitions, referring to the corresponding methodological sections, also depicting the numbers of records issued after the various filtering phases, would be of help;
- Section 5.1, page 18, caption of Fig. 5: since illustrations have to be self-explanatory, please explain the abbreviation “IQR” used in the figure legend; I would also recommend not using the abbreviation CV in the caption but rather the full word;
- Section 5.4, page 24, caption of Fig. 10: similarly, I would recommend not using the abbreviation CV in the caption but rather the full word;
- Section 5.4, page 24, lines 555-556: please consider rephrasing as follows, using the plural: “..., meaning that, for the majority of the test locations, the predicted probability scores for the respective landslide observations were higher than the predicted probabilities for any pre-landslide absence observations.”
- Section 5.4, page 24 line 558: please remove the word “a” from “suggests a-slightly lower model performances” since you mean a plural term;

- Section 5.4, page 24 line 559: please insert "the" in two locations as follows: "The influence of the test sample size on the variation in estimated..."
- Section 5.4, page 25, caption of Fig. 11: same observation, but "CV" may be put into brackets after the first use of the entire word to explain the abbreviation used further in the caption as well as on the figures themselves;
- Section 5.4, page 25, caption of Fig. 11, line 565: please continue with lowercase letters after "In a)";
- Section 5.4, page 26, caption of Fig. 12: similar observation as for the latter figure captions, regarding the use of the abbreviation "CV";
- Section 5.4, page 27, caption of Fig. 12, line 588: I would suggest reversing the sentence into: "A map of the environmental units is shown in Fig. 3", since this puts the emphasis on the map to ease understanding;
- Section 5.4, page 27, lines 596-597: please consider reversing the phrase as follows and adding "respectively" in the end: "The AUROC equalled 0.85 for the 7 shallow landslide locations and 0.9 for the 12 precipitation-induced flow-type landslides, respectively.";
- Page 31, line 740: please change the number of this section title to "7 Conclusions"
- I would suggest being consistent throughout the paper when writing the name of the model ("generalized additive mixed models"), i.e. with either upper- or lowercase letters.