Overview

The manuscript (MS) deals with the development of a mathematical tool useful for setting up a landslide early warning system (LEWS) in the Southern Andes, Chile. The methodology combines bias correction of precipitation products and a model for estimating the probability of landslide triggering. The topic is within NHESS and the Special Issue. Language and structure of the MS is acceptable but should be improved. I think the MS has potential for publication, but the current version needs major revisions, mainly because some important methodological aspects need to be explained more clearly and thus it is difficult to understand how scientifically sound are the results.

A: We appreciate your evaluation of our manuscript. We assessed all the comments to increase the quality of the MS. Following, we explain how the reviews were corrected/added.

Specific comments

Abstract needs to be improved: the various sentences are not adequately linked, so it is difficult to understand what is being done in the MS

A: We agree with the reviewer comment. We addressed this point in previous comment. Thanks you.

Methodology: it should be better explained how the results are supposed to be used within a LEWS. How is rainfall supposed to be used as input to the developed models to produce a warning? Which is the value of probability for which a warning should be issued?

A: Thanks you for this observation. Now, introduce additional information respect to the precipitation and geomorphological features.

Original text

Precipitation data and local geomorphological features were integrated into a logistic model to evaluate the occurrence of RIL.

Modified text:

Precipitation data and local geomorphological features were integrated into a logistic model as predictors to evaluate the occurrence of RIL.

Moreover, now improved the text including additional explanations about how the results will be used in LEWS. Original text:

None

Modified text:

We consider an approximation of the probability of occurrence of RIL through logistic distributions. The probability allows a spatialization of "prone-landslide" or "not prone-landslide" conditions under established precipitation and topographical conditions.

Regarding the probability for which a warning should be issued, now we introduce

Added text:

We propose that the threshold must be suitable to separate a prone-landslide event from a non-prone-landslide event. The threshold maximizes the sensibility in the four models with different degrees of performance of RILEWS.

Section 3.3: When computing ROC sensitivity and specificity how do you treat observed landslides? I mean, observed landslides are point features, while the output of your model is spatially distributed: how is the comparison between the two done? Is a buffer considered around observed landslide points, or you just take the value at the cell including the point?

A: Thanks you for this observation. Now, we clarify the comparison method

Original text

The quality of each regression was evaluated by ROC analysis (Fawcett, 2006) using the independent database BD2 (Figure 2).

Modified text:

The quality of each regression was evaluated by ROC analysis (Fawcett, 2006) using the independent database BD2 (Figure 2). The DB2 has georeferenced the initial failure zone. We compared the initial failure zone to the pixel of our models (pixel that includes the point).

Figure 8 and 9, model 1 and model 4 have basically the same performance. This means that Seven-day precipitation does not add much information. Perhaps the authors should think and comment on this

A: Thanks for your accurate observation. We agree. Now we modified the text:

This is consistent when we compare M1 (AUC=0.80) and M4 (AUC=0.79); they present similar sensitivity values (~91% in both cases), suggesting that either model could be used.

Added text:

This is consistent when we compare M1 (AUC=0.80) and M4 (AUC=0.79); they present similar sensitivity values (~91% in both cases), suggesting that either model could be used. Model 1 and Model 4 showed similar performance due to one being contained in the second. Hence, we interpret that an overrepresentation could exist. Therefore, model 4 does not support additional information. At an operational level, the discard of model 4 reduces the computing loading, simplifying the alert processes.

Table 2 – Model 4 is the only one combining 3 explanatory variables. Why do not investigate also all the other possible combinations of 3 and 4 variables?

A: We agree that multiple combinations may contribute to accurate models. We noted that daily and 7-days accumulated precipitation are sufficient to create RILEWS. A RILEWS with many precipitation-derived variables will need additional computational time, increasing the load of the operational system. Therefore, we propose to use limited variables in the RILEWS proposed. We added additional information in section 3 (previous to figure 2).

Figure quality should be improved (all)

A: We agree with the reviewer. Now we upload vectorial format to the server, increasing the quality.

4.2 is mainly a list of the calibrated parameters for the logit and probit models. Perhaps revise but creating a table with the parameters' values, while the text comments the table

A: We appreciate your useful comment. Now, we added two tables to section 4.2 to improve the understood of our contribution. Thank you

Added tables:

TABLE XXX1: Values of the estimators for the Logit models

	Intercept	**Daily precipitation**	**Seven-day accumulated precipitation**	**Thirty-day accumulated precipitation**	**Slope**
Model 1	3.5235 ± 0.0069	- 0.8176 ± 0.0089	-	-	-0.1696 ± 0.0049
Model 2	3.3582±0.0067	-	0.6413 ± 0.0063	-	-0.1365 ± 0.0086
Model 3	3.1658±0.0091	-	-	-0.3518 ± 0.0033	-0.1289 ± 0.0072
Model 4	3.5206±0.0106	- 0.8124 ± 0.0066	0.0020 ± 0.0086	-	-0.1675 ± 0.0080

TABLE XXX2: Values of the estimators for the Probit models

	Intercept	**Daily precipitation**	**Seven-day accumulated precipitation**	**Thirty-day accumulate d precipitatio**	**Slope**

			n		
Model 1	1.9113 ± 0.003	- 0.4166 ± 0.0046	-	-	- 0.0741 ± 0.0022
Model 2	1.8490±0.0031	-	-0.3545 ± 0.0029	-	-0.0675 ± 0.0038
Model 3	1.7482±0.0041	-	-	-0.1897 ± 0.0020	- 0.0596 ± 0.0033
Model 4	1.9110±0.0044	- 0.4016 ± 0.0027	- 0.0202 ± 0.0038	-	- 0.0732 ± 0.0040

Section 5.1 Precipitation accuracy, and about uncertainty in general: the paper may benefit from a more complete literature overview on this point: see, in addition to cited papers e.g.: https://doi.org/10.1016/j.geomorph.2015.04.028 , https://doi.org/10.1016/J.GEOMORPH.2014.06.015 https://doi.org/10.5194/hess-21-4525-2017 https://doi.org/10.1016/j.jhydrol.2015.10.010 https://doi.org/10.1016/j.geomorph.2016.11.019 https://doi.org/10.1016/j.geomorph.2017.02.001 https://doi.org/10.1007/s11069-018-3508-4 https://doi.org/10.1007/s11069-019-03830-x

A: We agree. Now we added additional information about the precipitation accuracy. We appreciate the references provided that allow expanding the value of our contribution.

L 134: you use only 20 – 30 % of the data for calibration. Why not an higher percentage?

A: We select this amount to has a “large” database form validation. The amount of RIL in the zone still poor in comparison to others studies. However, this correspond to the more dense database in the South of Chile that exists. Therefore, we want to develop a coarse logistic models using the worst data constrains.

It is unclear how you select 57 landslide events from the available 4,987 RIL. 57 events are a quite few, according to the literature (see, e.g. DOI: 1007/s10346-021-01704-7; https://nhess.copernicus.org/articles/21/2125/2021/)

A: Currently, the south of Chile has a well-identified landslide database without a date. The local database has 4987 RIL (with and without date). We select the RIL events inside the study zone that has precise dates. In to future, we expect to delimitate the landslide dates using the methodology of Morales et al. (2021). This constraint will allow us to develop better RILEWS and increase the spatial coverage of our system. Now, we correct the sentences:

Original text:

However, we had 2,035 of these, including the exact date. We used as database 57 RIL events, considering mudflow, debris flow and mass wasting.

Modified text:

However, 2.035 RIL exist in the zone, and only 57 RIL events have an exact date. The final database considers mudflow, debris flow and mass wasting. The actual database is not suitable to establish RILEWS using thresholds due to the scarce amount (Peres and Cancelliere, 2021). The current dataset is the most comprehensive landslide catalogue for the zone in comparison to well-validated global datasets such as Global Landslide Catalog (GLC) (Kirschbaum et al., 2010, 2015) and the Global Fatal Landslide Database (GFLD) (Froude and Petley, 2018) developed into other studies (Destro et al., 2017; Rossi et al., 2017; Wang et al., 2021).

References:

Froude, M. J. and Petley, D. N.: Global fatal landslide occurrence from 2004 to 2016, Nat. Hazards Earth Syst. Sci., 18, 2161–2181, https://doi.org/10.5194/nhess-18-2161-2018, 2018. a, b, c

Destro, E., Marra, F., Nikolopoulos, E. I., Zoccatelli, D., Creutin, J. D., and Borga, M.: Spatial estimation of debris flows-triggering rainfall and its dependence on rainfall return period, Geomorphology, 278, 269–279, https://doi.org/10.1016/j.geomorph.2016.11.019, 2017.

Kirschbaum, D. B., Adler, R., Hong, Y., Hill, S., and Lerner-Lam, A.: A global landslide catalog for hazard applications: method, results, and limitations, Nat. Hazards, 52, 561–575, 2010. a, b

Peres, D. J. and Cancelliere, A.: Comparing methods for determining landslide early warning thresholds: potential use of non-triggering rainfall for locations with scarce landslide data availability, Landslides, 18, 3135–3147, https://doi.org/10.1007/s10346-021-01704-7, 2021.

Rossi, M., Luciani, S., Valigi, D., Kirschbaum, D., Brunetti, M. T., Peruccacci, S., and Guzzetti, F.: Statistical approaches for the definition of landslide rainfall thresholds and their uncertainty using rain gauge and satellite data, Geomorphology, 285, 16–27, https://doi.org/10.1016/j.geomorph.2017.02.001, 2017.

Wang, X., Otto, M., and Scherer, D.: Atmospheric triggering conditions and climatic disposition of landslides in Kyrgyzstan and Tajikistan at the beginning of the 21st century, Nat. Hazards Earth Syst. Sci., 21, 2125–2144, https://doi.org/10.5194/nhess-21-2125-2021, 2021.

Technical corrections

I have annotated the manuscript with some technical corrections (See attachment).
Citation: https://doi.org/10.5194/nhess-2021-317-RC2

A: We reviewed the technical corrections. We rewrote the sentences/words that did not suffer changes due to the previous comments