GENERAL COMMENTS

The contribution "Evaluation of predictive models for post-fire debris flows occurrence in the western United States" by Efthymios I. Nikolopoulos and co-Authors is good and potentially publishable. The Authors in this study we investigate the potential to improve the efficiency of current predictive models with machine-learning approaches. The analysis is based on a database on post-fire debris flows published by United States Geological Survey. In general, the topic well address scientific questions within the scope of NHESS.

The theoretical background is well-argued. Review of literature seems complete. The description of study area is not sufficiently complete. The description of methodology and successive parts of paper are well organized but can be improved. Results, and discussion sections are short compared to amount of work done. They should be widely increased. The readability of the whole paper is sufficient with a quite good English. Overall, the work presents some carelessness and incompleteness. It can be published on NHESS journal only after a major revision.

SPECIFIC COMMENTS

I have some specific comments that should be addressed before the manuscript can be accepted for publication.

- Section 1, "Introduction" (Page 2-3).

At Page 3, Lines 23 - 25 the Authors should add references about the rainfall accumulation-duration thresholds, and the two model used.

- Section 2, "Study area and data" (Page 4, Lines 1-30).

This paragraph is confuse and not complete. The Fig.1 does not show well the study area, the four different regions, the location of the DF, and the area of fire-affected catchments. In addition, the Authors should better explain the differenced normalized burn ratio (dNBR) used in your analysis (Page 4, Lines 13-14).
For a better clarity and understanding of the text, Authors should specify in detail how the values of the average erodibility index (KF-Factor) for the four different regions were chosen (Page 4, Lines 16-21).

- Section 3, "Seasonality and characteristics of rainfall events" (Page 5-7).

At the page 5, Lines 16-24 the Authors refer to statistics about the seasonality and characteristics of rainfall events reported in Table 2. The total number of the events is normalized in terms of percentage, this well describe the distribution of rainfall event analyzed for the different regions. However, this does not allow a comparison by number of events for each region. This comparison could be useful to guarantee the correctness of the considerations set out in Section 3 (Page 6, lines 4-28) and in Section 4.1.3 (Page 9, lines 6-8). I suggest to integrate in the Table 2 the absolute and relative frequency of the DF and noDf events.

- Section 4.1.2 "Logistic regression" (Page 7-8).

The LR model works with either continuous or categorical independent variables, or a combination of the two types, regardless whether they present a normal distribution or not (Costanzo et al., 2014). The underlying mathematical relationship between the dependent dichotomous variable (presence/absence of a landslide in the mapping unit and the n independent variables (X_1, ...,X_n). Authors adopt the logistic regression (LR) model using a set of explanatory variables (X_1, X_2, and X_3) that have in common the cumulated rainfall (max 15 min rainfall accumulation). My question is: the explanatory variables used for the Authors are really independent? For a better clarity and understanding, Authors should specify in detail this aspect.

- Section 4.1.3 "Random forest" (Page 8-9).

Authors stated that "*The first model (RF-ED) was developed using the variables of rainfall accumulation and duration. It is the model that we consider as the one with minimum data requirements, given that only two rainfall variables are used for the prediction.*" (Page 9, Lines 10-12), but in the Page 9, (Lines 2-5) the Authors indicate they have used also an extra categorical variable (named "Region class"). The authors should better explain the total number of variables used and verify if it is in agreement with Table 3.

- Section 4.3 "Identification of thresholds" (Page 10, Lines 11-25).

This section is very short. Authors should better argue and comment the obtained results. I suggest to include in the paper a new Figure that show the distribution of the (D,E) pairs with the rainfall thresholds and their uncertainties. In addition the Authors should better explain how they reconstructed the rainfall conditions responsible of the DF and noDF. In particular, if have been used the subjective or objective methods (Vessia et al., 2014; Rossi et al. 2017; Melillo et al. 2015, 2018).

TECHNICAL CORRECTIONS

Page 1, Line 21: After Page 1, Line 20 it seems that there are two parts in the text that are not well connected. I suggest to rewrite better the start of this part.

Page 1, Line 24: I suggest to introduce here the acronyms of the four regions (AZ, CA,CO, and NM), because are here that are cited for the first time.

Page 7, Line 21: I suggest to replace "both debris and no debris flow events" with "both DF and noDF events"

Page 8, Line 1: The acronym "PFDF" does not exist. I suggest to introduce PFDF in the previous part (Page7, Line 7)

Page 8, Line 12, 14, 15: I suggest to replace "X_{1}, X_{2}, X_{3}" with "X_1, X_2, X_3".

Page 10, Line 18, 22: I suggest to replace "NO DF" with "noDF".

Page 11, Line 7: I suggest to replace "NO DF" with "noDF".

Page 19, Table 1: I suggest to replace "(DF=1,NO DF=0)" with "(DF=1, noDF=0)"

Page 22, Table 1: I suggest to replace "(NO Debris Flow)" with "no Debris Flow" and "(No Debris Flow)" with "no Debris Flow".

Page 26, Figure 1: Please show better in the Figure the study area, the four different regions, the location of the DF, and the area of fire-affected catchments. In addition, insert horizontal scale and North indicator symbol.

Page 27, Figure 2: In the blue (AZ region) and black (CO region) boxplots the median is not well visible. Please change the colours using other with major contrast. In addition I suggest to replace in the caption of the Figure "Figure 2 Boxplot for storm duration, storm accumulation and peak 15-min storm accumulation" with "Figure 2 Boxplot for (a) storm duration, (b) storm accumulation and (c) peak 15-min storm accumulation".

Page 28, Figure 3: I suggest to replace the caption with "(a) Total rainfall accumulation vs duration and (b) peak 15 minute rainfall accumulation vs duration for Arizona (AZ), California (CA), Colorado (CO) and New Mexico (NM). Colored dots and x symbols correspond to DF and noDF occurrence respectively".

Page 30, Figure 5: I suggest to use a different type of the line (solid line) to represent the full range of variation. The dashed-line produce confusing.

References:

Costanzo, D., Chacón, J., Conoscenti, C., Irigaray, C., and Rotigliano, E. (2018). Forward logistic regression for earth-flow landslide susceptibility assessment in the Platani river basin (southern Sicily, Italy), Landslides, 11, 639–653, https://doi.org/10.1007/s10346-013-0415-3, 2014.

Melillo, M., Brunetti, M.T., Peruccacci, S., Gariano, S.L., Guzzetti, F., (2015). An algorithm for the objective reconstruction of rainfall events responsible for landslides. Landslides 12(2), 311–320, https://doi.org/10.1007/s10346-014-0471-3

Melillo M., Brunetti M. T., Peruccacci S, Gariano S. L., Roccati A., Guzzetti F. (2018). A tool for the automatic calculation of rainfall thresholds for landslide occurrence. Environmental Modelling & Software, https://doi.org/10.1016/j.envsoft.2018.03.024

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

Vessia, G., Parise, M., Brunetti, M. T., Peruccacci, S., Rossi, M., Vennari, C., Guzzetti, F., 2014. Automated reconstruction of rainfall events responsible for shallow landslides. Nat. Hazards Earth Syst. Sci.14(9), 2399-2408, https://doi.org/10.5194/nhess-14-2399-2014.