Anonymous Referee #1

The manuscript "Setting up the critical rainfall line for debris flows via support vector machines" by Y. F. Tsai, C. H. Chan, and C. H. Chang aims at improving debris-flow mitigation through critical rainfall line approach. This methods requires 1-classifying debris-flow prone zones into clusters and 2-attribute one critical rainfall line to each group through statistical analysis. This could yield to global, near-real time monitoring systems based only on rainfall data that may help authorities to deal with debris flow hazard.

I found the main idea of the manuscript very interesting, with a great potential for implementation in hazard mitigation systems. It does not require any underlying physical model, which allows quick and efficient cover at the country-scale.

However, I think that this work should benefit from significant improvement before publication in NHESS. I first list a series of global comments in the "Scientific Quality" section; then I move on to more technical aspects in the "Presentation Quality" section.

1-Scientific Quality:

1.1-The overall paper lacks of reference to existing work in the literature. This is critical in the "Introduction" section, with only two references listed. Only one refers to existing work from other research groups. I would advise to improve the introduction section with more precise and reference-supported steps (see a few papers in 2-Presentation Quality).

Thank you for your kind reminding. We have added a series references as follows.

Lin C.W., Shieh C.-L., Yuan B.D., Shieh Y.C., Liu S.H., Lee S.Y.: Impact of Chi-Chi earthquake on the occurrence of landslides and debris flows: example from the Chenyulan River watershed, Nantou, Taiwan, Engineering Geology, 71, 1–2, 49-61, 2004.

Liu S.H., Lin C.W., Tseng C.M.: A statistical model for the impact of the 1999 Chi-Chi earthquake on the subsequent rainfall-induced landslides, Engineering Geology, 156, 11-19, 2013.

Huang R., Li W.: Post-earthquake landsliding and long-term impacts in the Wenchuan earthquake area, China, Engineering Geology, 182, 111-120, 2014.

Lin C.W., Liu S.H., Lee S.Y., Liu C.C.: Impacts of the Chi-Chi earthquake on subsequent rainfall-induced landslides in central Taiwan, Engineering Geology, 86, 2–3, 87-101, 2006.

Chiou, S. J., Cheng, C. T., Hsu, S. M., Lin, Y. H., & Chi, S. Y.: Evaluating landslides and sediment yields induced by the Chi-Chi Earthquake and followed heavy rainfalls along the Ta-Chia River, Journal of GeoEngineering, 2, 2, 73-82, 2007.

Lin M.L., Tung C.C.: A GIS-based potential analysis of the landslides induced by the Chi-Chi earthquake, Engineering Geology, 71, 1–2, 63-77, 2004

Zhuang J.Q., Cui P., Wang G.H., Chen X.Q., Iqbal J., Guo X.J.: Rainfall thresholds for the occurrence of debris flows in the Jiangjia Gully, Yunnan Province, China, Engineering Geology, 195, 335-346, 2015.

1.2-The methodology of the study is incomplete. There is no reference to the rainfall data origin. The landslide classification method is unclear. What is the method for debris-flow detection after a rainfall event? Are the data GIS-based? On which basis? There is no discussion about the quality of these input data into the SVM model. Please provide information/ reference/ discussion supporting that those 8 characteristics are able to explain the debris-flow susceptibility.

We miss to describe the origin of data. In this new version, we have add origin of data, including rainfall, geological data and hydrological data source as "*geographical information from central geological survey MOEA, hydrological information from central weather bureau in Taiwan, historical data of damage from internet-based news*" in PAGE 5 LINE 14-16. These data above are GIS-based. For 8 characteristics, we have added three references for choosing factor reason as follows.

Wan S., Lei T.C.: A knowledge-based decision support system to analyze the debris-flow problems at Chen-Yu-Lan River, Taiwan, Knowledge-Based Systems, 22, 8, 580-588, 2009.

Yuan L.F., Zhang Q.F., Li W.W., Zou L.J.: Debris Flow Hazard Assessment Based on Support Vector Machine, Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference, 4221-4224, 2006.

Bui D.T., Pradhan B., Lofman O., Revhang I.: Landslide Susceptibility Assessment in Vietnam Using Support Vector Machines, Decision Tree, and Naïve Bayes Models, Mathematical Problems in Engineering, 1-26, 2012.

1.3-The FCGA method: how was set the number of stream groups? (7). Quantify the semblance/difference between groups using metrics and discuss the results.

FCGA clustering and hierarchical tree results shows that the principle during clustering is referred to the patterns (rock type (R), watershed area (A), effective watershed area (A_{15}), landslide area (A_s), landslide ratio(A_s/A), length of channel in the effective watershed area (L), mean surface slope of the effective watershed area (S_s) and mean channel slope of the effective watershed area (S_c)) of debris flow streams. We have added a new figure which is the results of family competition genetic algorithm and hierarchical tree analysis (Fig. 2 and Fig. 3), the seven groups are built on the basis of this results. Moreover, this result is computed optimally and generated automatically by program in PAGE 4 LINE 26 and PAGE 5 LINE 1-5.

1.4- The SVM method: please provide metrics for critical rainfall line choice. The rainfall lines choice for each group (Fig. 7) is not clear nor unambiguous. Please explain the method and discuss the results.

Thank you for your reminding. About the resolution of Fig. 7, we have redrawn those figures in clear color. Due to lack of damage occurrence data of the first four years, this study adopts this method, support vector machine, can define the maximum hyper plane (also critical rainfall line) under few data point without overfitting problem.

1.5- Discussion/conclusions: The work presented here is mostly qualitative. It requires quantitative result presentation and discussion to support the feasibility of SVM-based critical rainfall line setting. This SVM approach for critical rainfall line has been widely applied to debris-flow prediction (see references provided). To my opinion, the paper has to focus more on its originality and strength: the application of this technique in Taiwan during the post- Chi-Chi earthquake period.

This study is focused on the critical rainfall thresholds changed after 1999 921 earthquake in Taiwan, hence, this study collects this most effected four years data after earthquake to comprehend the variation of critical rainfall thresholds. However, there is not enough debris flow occurrence data of each debris flow streams. Therefore, this study utilizes FCGA to combine different debris flow streams which have the same significant patterns, and this procedure can also increase the accuracy of analysis results because the numbers of debris flow occurrence data have been raised. At last, this study adopts support vector machine to estimate the critical rainfall threshold of each group by their significant factors (Discussion/Conclusion).

2- Presentation Quality:

This section mostly deals with english/typo/clarity of the text.

Propositions for modifications are written in *italic*.

PAGE 5958:

L2. "to help preventing"

Thank you for your kind reminding. We have modified this sentence in PAGE 1 LINE 13.

L6. please precise "similar" in a quantitative manner.

The term "*similar*" is based on FCGA clustering and hierarchical tree results because the principle during clustering is referred to the patterns (rock type (R), watershed area (A), effective watershed area (A_{15}), landslide area (A_s), landslide ratio(A_s/A), length of channel in the effective watershed area (L), mean surface slope of the effective watershed area (S_s) and mean channel slope of the effective watershed area (S_c)) of debris flow streams.

L11. Please precise the meaning of "structural risk"

Thank you for your reminding, we replace the wrong word "structural risk" to "disaster risk", which represents the potential of disaster and its sensitivity PAGE 1 LINE 27.

L17. Please check "yearly often"

Thank you for your reminding. We have modified the words "*yearly often*" as "*every year*" because typhoon and Mei-Yu weather phenomenon appear in Taiwan every year PAGE 2 LINE 4.

L19. " Especially after the Chi-Chi earthquake...in the center of Taiwan". Please provide reference.

Thank you for your reminding. We have added three references for the comments as follows.

Lin C.W., Shieh C.-L., Yuan B.D., Shieh Y.C., Liu S.H., Lee S.Y.: Impact of Chi-Chi earthquake on the occurrence of landslides and debris flows: example from the Chenyulan River watershed, Nantou, Taiwan, Engineering Geology, 71, 1–2, 49-61, 2004.

Liu S.H., Lin C.W., Tseng C.M.: A statistical model for the impact of the 1999 Chi-Chi earthquake on the subsequent rainfall-induced landslides, Engineering Geology, 156, 11-19, 2013.

Huang R., Li W.: Post-earthquake landsliding and long-term impacts in the Wenchuan earthquake area, China, Engineering Geology, 182, 111-120, 2014.

L19. "These tremendous landslides often brought sediment material into the streambed, in the initiation area of debris flow".

Thank you for your kind reminding. We have modified as his sentence in PAGE 2 LINE 5-6.

L21. "will be mobilized by rainfall"

Thank you for your kind reminding. We have modified this sentence in PAGE 2 LINE 7.

L24. "will have". Please provide reference.

Thank you for your reminding. We have added two references for the comments as follows.

Lin C.W., Liu S.H., Lee S.Y., Liu C.C.: Impacts of the Chi-Chi earthquake on subsequent rainfall-induced landslides in central Taiwan, Engineering Geology, 86, 2–3, 87-101, 2006.

Chiou, S. J., Cheng, C. T., Hsu, S. M., Lin, Y. H., & Chi, S. Y.: Evaluating landslides and sediment yields induced by the Chi-Chi Earthquake and followed heavy rainfalls along the Ta-

Chia River, Journal of GeoEngineering, 2, 2, 73-82, 2007.

L25. " This means that the debris flow disasters have been more unpredictable and destructive with the amount of sediment material". Please provide reference.

Thank you for your reminding. We have added one reference for the comments as follows.

Lin M.L., Tung C.C.: A GIS-based potential analysis of the landslides induced by the Chi-Chi earthquake, Engineering Geology, 71, 1–2, 63-77, 2004

L26. "The numerous landslides triggered .... (Shieh et al. 2009)."

Thank you for your kind reminding. We have modified this sentence in PAGE LINE.

PAGE 5958:

L2. "Lowered their .."

Thank you for your kind reminding. We have modified this sentence in PAGE 2 LINE 12.

L3. " Nakamura et al. (2000) also reported a huge... earthquake in Japan."

Thank you for your kind reminding. We have modified this sentence in PAGE 2 LINE 14-16.

L4. "Almost every landslide during that time induced server debris-flow disasters (Fig. 1)". Please provide reference.

Thank you for your reminding. We have deleted this sentence because we remove figure 1.

L6. "Thus, in order to prevent the disaster of debris flow". Please precise the prevention method underlying.

Thank you for your reminding. Through establishing more accurate critical rainfall line for debris flow streams can efficiently alert beforehand, therefore, it is important to define critical rainfall line of each debris flow streams.

L-. "We aim to set"

Thank you for your kind reminding. We have modified this sentence in PAGE 2 LINE 19.

L8-9. Unclear sentence. "Firstly, 377 debris-flow streams in the center of Taiwan affected by Chi-Chi earthquake were considered." Please precise "affected".

Thank you for your kind reminding. Large earthquakes might trigger landslides appearance, we can refer to these papers as follows.

Lin C.W., Shieh C.-L., Yuan B.D., Shieh Y.C., Liu S.H., Lee S.Y.: Impact of Chi-Chi earthquake on the occurrence of landslides and debris flows: example from the Chenyulan River watershed, Nantou, Taiwan, Engineering Geology, 71, 1–2, 49-61, 2004.

Liu S.H., Lin C.W., Tseng C.M.: A statistical model for the impact of the 1999 Chi-Chi earthquake on the subsequent rainfall-induced landslides, Engineering Geology, 156, 11-19, 2013.

Huang R., Li W.: Post-earthquake landsliding and long-term impacts in the Wenchuan earthquake area, China, Engineering Geology, 182, 111-120, 2014.

L11. "Streams with similar characteristics were then clustered together and support vector machines (SVMs) were applied to setup the critical rainfall line for each debris-flow cluster. "

Thank you for your kind reminding. We have modified this sentence in PAGE 2 LINE 23-25.

L13. "The experimental results show that SVM method performs well in setting a critical rainfall line for each group of debris-flow."

Thank you for your kind reminding. We have modified this sentence in PAGE 2 LINE 25-26.

L17. "... Taiwan and caused..."

Thank you for your kind reminding. We have modified this sentence in PAGE 2 LINE 29.

L18. " the blocks of these landslides were up to 2365 units". Please precise "blocks" and "units"

Thank you for your kind reminding. We found that word "block" is mistake, so we have modified this word as "" in PAGE 2 LINE 30.

L19. "are mostly located"

Thank you for your kind reminding. We have modified this sentence in PAGE 2 LINE 31.

L20. Unclear sentence. "This study focuses on the debris flow streams originating from landslides affected by the Chi-Chi earthquake".

Thank you for your reminding. We have added one reference for the comments as follows in PAGE 3 LINE 1.

Yu B., Wu Y., Chu S.M.: Preliminary study of the effect of earthquakes on the rainfall threshold of debris flows, Engineering Geology, 182, 130-135, 2014

L24. (Fig. 2b).

Thank you for your kind reminding. We have modified this sentence in PAGE 3 LINE 3-4.

PAGE 5960:

L5. " In order to cluster the 377 debris flow streams into different groups" Please explain why different groups are needed for further analysis.

This study is focused on the critical rainfall thresholds changed after 1999 921 earthquake in Taiwan, hence, this study collects this most effected four years data after earthquake to comprehend the variation of critical rainfall thresholds. However, there is not enough debris flow occurrence data of each debris flow streams. Therefore, this study utilizes FCGA to combine different debris flow streams which have the same significant patterns, and this procedure can also increase the accuracy of analysis results because the numbers of debris flow occurrence data have been raised in PAGE 3 LINE14-18.

L6. Please precise the origin of geographical information, hydrological information, historical data of disaster and statistical tables.

We miss to describe the origin of data. In this new version, we have add origin of data, including rainfall, geological data and hydrological data source as "*geographical information from central geological survey MOEA, hydrological information from central weather bureau in Taiwan, historical data of damage from internet-based news*" in PAGE 3 LINE 14-18.

L7. "have to be .."

Thank you for your kind reminding. We have modified this sentence in PAGE 3 LINE 18.

L13. "each stream"

Thank you for your kind reminding. We have modified this sentence in PAGE 3 LINE 22.

L17. "After data normalization, ... can be ... "

Thank you for your kind reminding. We have modified this sentence in PAGE 3 LINE 27.

L25. Please precise the meaning of "structural risk"

Thank you for your kind reminding. We have modified this sentence in PAGE 5 LINE 22.

L27. "be performed"

Thank you for your kind reminding. We have modified this sentence in PAGE 5 LINE 24.

PAGE 5961:

L4. " Since this term measures distance"

Thank you for your kind reminding. We have modified this sentence in PAGE 4 LINE 27.

L7. " This section aims to cluster 377 debris-flow streams into seven groups, via clustering analysis such that streams in each group have similar characteristics." Please precise meaning of "similar" in a quantitative manner.

The term "*similar*" is based on FCGA clustering and hierarchical tree results, and the principle of these methods during clustering is referred to the patterns (rock type (R), watershed area (A), effective watershed area (A_{15}), landslide area (A_s), landslide ratio(A_s/A), length of channel in the effective watershed area (L), mean surface slope of the effective watershed area (S_s) and mean channel slope of the effective watershed area (S_c)) of debris flow streams. These methods can compute the most similar patterns of debris flow streams together by calculating their distances.

L8. " An efficient clustering algorithm". Please provide proof/reference/metrics of this algorithm and its efficiency.

Thank you for your kind reminding. We have added one reference for this aspect as follows in PAGE 4 LINE 15-17.

Yang J.M., Kao C.Y.: A family competition evolutionary algorithm for automated docking of flexible ligands to proteins, 4, 3, 225-237, 2000.

L10. " This approach was employed to group 377 debris flow streams into seven groups". Please explain why seven? Provide metrics or reference.

Thank you for your kind reminding. We have added a new figure which is the results of family competition genetic algorithm and hierarchical tree analysis, the seven groups are built on the basis of this results. Moreover, this result is computed optimally and generated automatically by program in PAGE 4 LINE 17-19.

L19. "In this study, a family competition genetic algorithm (FCGA) was used to construct a hierarchical tree of streams."

Thank you for your kind reminding. We have modified this sentence in PAGE 4 LINE 25-26.

PAGE 5962:

L1. " The concepts of family competition have been successfully applied to solve numerous continuous parameter optimization problems, including protein docking (Yang and Kao, 2000). In the authors' earlier work (Tsai et al., 2001), family competition and EAX were successfully integrated to solve traveling salesman problems (TCPs). Neighbor-join mutation (Tsai et al., 2002) was developed to coordinate with the EAX and thus balance exploration and exploitation." This paragraph is not directly relevant to the issue of this manuscript. Please cite only appropriate references such as "Further detail about family competition and efficiency can be found in .... + references". Please complete the references with relevant works from other research groups than the authors one.

Thank you for your kind reminding. We have accepted referee's advice, so this paragraph has been removed.

L6. "The method used in this study combines family competition , Neighbor-join mutation and EAX."

Thank you for your kind reminding. But this paragraph has been removed.

L9. " The experimental results revealed that the FCGA is a promising method for constructing the optimal tree of streams." Please present the results first. Then conclude on the FCGA efficiency.

Thank you for your kind reminding. We have added new figure of clustering results in Fig. 2 in PAGE 5 LINE 4.

L11. " Figure 3 presents the seven groups of 377 debris flow streams." Please provide metrics on the figure/text (distance between groups, etc).

Thank you for your kind reminding. We have added a new figure which is the results of family competition genetic algorithm and hierarchical tree analysis in PAGE 5 LINE 4.

L12. "each group"

Thank you for your kind reminding. We have modified this sentence in PAGE 5 LINE 9.

L13. "Each groups all exhibited different trends in their characteristics and the characteristics in the same group were similar." Please provide quantitative information to support that statement.

This decription is based on FCGA clustering and hierarchical tree results, and the principle of these methods during clustering is referred to the patterns (rock type (R), watershed area (A), effective watershed area (A_{15}), landslide area (A_s), landslide ratio(A_s/A), length of channel in the effective watershed area (L), mean surface slope of the effective watershed area (S_s) and mean channel slope of the effective watershed area (S_c)) of debris flow streams. These methods can compute the most similar patterns of debris flow streams together by calculating their distances in Fig. 2.

L17 " a possible mean"

Thank you for your kind reminding. We have modified this sentence in PAGE 5 LINE 13.

L17. " As a result, this method represents a possible means of establishing a critical rainfall line for debris flow streams in each group". Please precise the meaning/underlying assumption.

Thank you for your kind reminding. Because every debris flow streams in each group have the same patterns inside (shows in Fig. 2 and Fig. 3), as the result, the critical rainfall line of each group can represent or reflect the distinguish patterns of themselves.

L18-19. " The critical rainfall lines of each groups could be set according to the characteristics." Please detail.

Thank you for your kind reminding. The clustering methods, FCGA and hierarchical tree, are analyzed on basis of the attribute conditions of each debris flow streams in PAGE 5 LINE 14-15.

L25. Please precise "structural risk"

Thank you for your kind reminding. We have replaced the word “structural risk” to “risk” only in PAGE 5 LINE22.

L27 "be performed"

Thank you for your kind reminding. We have modified this sentence in PAGE 5 LINE 24.

PAGE 5963:

L3. "This study intends to establish the critical rainfall line of debris flow via SVM."

Thank you for your kind reminding. We have modified this sentence in PAGE 5 LINE 28.

L5. "In this study" repetition of study two times

Thank you for your kind reminding. We have modified this sentence as “*Each data of debris flow stream were consider as a vector or a point in a multidimensional space.*” in PAGE 5 LINE 29-31.

L6. " The hyper-plane separating the vectors into two parts is then searched for, according to the occurrence of debris flow(Fig. 4)".

Thank you for your kind reminding. We have modified this sentence in PAGE 5 LINE 30-31.

L9. " However, two problems are frequently encountered during the process of classification."

Thank you for your kind reminding. We have modified this sentence in PAGE 6 LINE 2-4.

L10. Please rephrase in a clearer manner. " Figure 5 shows the two problems, it is likely that there are many hyper-planes existed in the multidimensional space, or it does not have any hyper-plane could separate the training data into two parts exactly."

Thank you for your reminding. We have modified this paragraph as follows, “*However, two problems are frequently encountered during the process of classification. Figure 5 shows the two problems, it is possible that there are many hyper-planes existed in the multidimensional space exactly. (Tax and Duin, 2002)*”.

L13-25. Please provide references and rewrite this paragraph in a more concise way.

Thank you for your kind reminding. We have added two references and modified the paragraph as follows in PAGE 6 LINE6-25.

Vapnik, V.: Statistical learning theory*. New York: John Wiley & Sons, 1998.*

Ballabio C., Sterlacchini S.: Support Vector Machines for Landslide Susceptibility Mapping: The Staffora River Basin Case Study, Italy, Mathematical Geosciences, 44, 1, 47-70, 2012.

PAGE 5964:

L16. Please discuss the risk of "false alarms" with this classification. And the risk of debris flow without alarm. How to balance the two issues?

Thank you for your kind reminding. This paper is focused on establishing critical rainfall lines and trying to figure out their changes after large earthquake in a short period, as the result, this paper dose not discuss about the false alarm problems.

PAGE 5965:

Nakamura et al. (2000) reference is missing.

Thank you for your kind reminding. We have modified this missing.

PAGE 5970:

Figure 1 please improve resolutions/provide vector.

Thank you for your kind reminding. We have removed this figure.

PAGE 5975:

Figure 7. Please improve the legend, which is vague. Please change the dot size on figures. Please change blue point to another color, more visible on the green area.

Thank you for your kind reminding. We have modified these figure in more visible color to instead original ones.

Please consider literature, such as (non-exhaustive):

Thank you for your kind suggestion. We have added all papers you suggested as follows.

Liang W.J., Zhuang D.F., Jiang D., Pan J.J., Ren H.Y.: Assessment of debris flow hazards using a Bayesian Network, Geomorphology, 171–172, 94-100, 2012.

X. Yao, Tham L.G., Dai F.C.: Landslide susceptibility mapping based on Support Vector Machine: A case study on natural slopes of Hong Kong, China, Geomorphology, 101, 4, 572-582, 2008.

Zhou, W., and Tang, C.: Rainfall thresholds for debris flow initiation in the Wenchuan earthquake-stricken area, southwestern China, Landslides, 11, 5, 877-887, 2014.

Zhuang J.Q., Cui P., Wang G.H., Chen X.Q., Iqbal J., Guo X.J.: Rainfall thresholds for the occurrence of debris flows in the Jiangjia Gully, Yunnan Province, China, Engineering Geology, 195, 335-346, 2015.