

***Author Responses Addressing Review from Referee #1 for “Transfer learning for landslide susceptibility modelling using domain adaptation and case-based reasoning” by Wang et al.***

For these responses, we address each Referee comment (RC) individually and include our response below it. The Referee Comments are numbered and use a black font, while the Author Responses (AR) use a red font.

Thank you for giving me the opportunity to revise the manuscript entitled “Transfer learning for landslide susceptibility modelling using domain adaptation and case-based reasoning” by Zhihao Wang and his/her colleagues that was submitted to “Geoscientific Model Development”. The manuscript used transfer-learning strategies to landslide susceptibility assessment (LAS) for 10 different cases in Austria, Ecuador, and Italy. Application of various learning based algorithms are interesting topics in LSA, but regarding the existed literatures needs to be polished properly.

AR0.0: We want to thank Referee #1 for the helpful comments and suggestions. We have done our best to address each of the comments below.

RC1. The English of the paper is readable; however, I would suggest the authors to have it checked preferably by a native English-speaking person to avoid any mistakes.

AR1: Thanks for the comment and sorry for some mistakes. One of co-authors Jason Goetz is a native speaker, who is from Canada, and we also took another careful look at the revised manuscript.

RC2. Add the “...several case studies” in the title.

AR2: The objective of our study is to evaluate performance of transfer learning in landslide assessments. Hence, we performed experiments to validate the methods in the manuscript based on our available data, rather than specifically targeting these areas.

We might also identify such similarities in the recommended cites below.

For example,

1. Fang, Z., Wang, Y., Peng, L., & Hong, H. (2020). Integration of convolutional neural network and conventional machine learning classifiers for landslide susceptibility mapping. *Computers & Geosciences*, 139, 104470.

Authors wanted to assess landslide susceptibility by integrating convolutional neural network (CNN) with three conventional machine learning classifiers in the case of Yongxin Country, China. Because they focus on assessing the performance of the method, the case study was used as experimental study area. Therefore, they didn't add case study on the title.

Also, the same is true for the two recommended citations below.

2. Fang, Z., Wang, Y., Peng, L., & Hong, H. (2021). A comparative study of heterogeneous ensemble-learning techniques for landslide susceptibility mapping. *International Journal of Geographical Information Science*, 35(2), 321-347.

3. Azarafza, M., Azarafza, M., Akgün, H., Atkinson, P. M., & Derakhshani, R. (2021). Deep learning-based landslide susceptibility mapping. *Scientific reports*, 11(1), 1-16.

Therefore, in order to better reflect the aim of our work, we may not recommend adding "...several case studies" in the title. In Discussion, we also added another paragraph in section 4.3:

"Although the study areas cover a wide range of climates with different land cover types and landslide process types, our set of source areas is by no means complete and the results may therefore not be fully representative for the performances that might be achieved at a global scale. Future work should therefore broaden the database of source areas."

RC3. The concluding remarks of the abstract are not well-written. It's merely the repetition of the objectives and title of the manuscript. In an appropriate abstract, quantitative finding, method limitations, model verification, learning rate and justification have to be added. Please kindly modify the abstract.

AR3: Thanks for the comment. According to the preparation of *Geoscientific Model Development (GMD)* journal, the description for abstract in the Manuscript composition part said:

"The abstract should be intelligible to the general reader without reference to the text. After a brief introduction of the topic, the summary recapitulates the key points of the article and mentions possible directions for prospective research."

For our abstract, the first sentence is to introduce our objective. The rest of it summarizes results and their broader implication, which we believe is appropriate for a Conclusion. The objective of our work is to compare the methods used and analyze their potential for landslide assessments. Hence, based on the GMD description - a concise summary of the key points of the article, we have concisely summarized the experimental results to highlight the purpose of our work.

Therefore, we believe that the abstract in the manuscript can be appropriate for our work.

RC4. Considering the existing literature, the motivation of the manuscript is not convincing. There is a large literature associated with LSA, and various techniques have been applied/developed to assess the susceptibility better. I expect to see a novelty in a manuscript dealing with landslide susceptibility because we already have many case studies. I do not think that new case studies would help the scientific community to go one step further. The result of this manuscript is only valid for the examined area. Based on site-specific conditions and the quality of a landslide inventory, different results can be obtained in another site, and another

technique may appear like the best alternative. However, in the end, these efforts do not help us decide on the best method of LSA. So, please kindly request to provide solid document to support the superiority of the method than other approaches. Please check and refer these papers:

1. Kavzoglu, T., Colkesen, I., & Sahin, E. K. (2019). Machine learning techniques in landslide susceptibility mapping: a survey and a case study. *Landslides: Theory, practice and modelling*, 283-301.

2. Saleem, N., Huq, M., Twumasi, N. Y. D., Javed, A., & Sajjad, A. (2019). Parameters derived from and/or used with digital elevation models (DEMs) for landslide susceptibility mapping and landslide risk assessment: a review. *ISPRS International Journal of Geo-Information*, 8(12), 545.

AR4: For our work, we focus on the performance of transfer learning in landslide assessments using a sample of numerous, diverse study areas worldwide. Therefore, it is not a case study.

Transfer learning is the application of knowledge gained from completing one task to help solve a different, but related, problem. In landslide assessments, transfer learning can use the existed landslide inventories to detect landslides in unseen areas. But it should notice that transfer learning techniques have not been systematically explored in the context of landslide susceptibility modeling. Nowadays, more and more research groups try to scale modeling effort up from the local scale to the regional or even global scale, e.g.:

1. Lima, P., Steger, S. & Glade, T. Counteracting flawed landslide data in statistically based landslide susceptibility modelling for very large areas: a national-scale assessment for Austria. *Landslides* 18, 3531–3546 (2021).

2. Lin, Q., Lima, P., Steger, S., Glade, T., Jiang, T., Zhang, J., ... & Wang, Y.. National-scale data-driven rainfall induced landslide susceptibility mapping for China by accounting for incomplete landslide data. *Geoscience Frontiers*, 12(6), 101248 (2021).

3. Martina Wilde, Andreas Günther, Paola Reichenbach, Jean-Philippe Malet & Javier Hervás Pan-European landslide susceptibility mapping: ELSUS Version 2, *Journal of Maps*, 14:2, 97-104 (2018).

4. Lin, L., Lin, Q., and Wang, Y.: Landslide susceptibility mapping on a global scale using the method of logistic regression, *Nat. Hazards Earth Syst. Sci.*, 17, 1411–1424 (2017).

Hence, it is of great importance to research scale modeling in landslide assessments. Our study is to study the landslide modeling in global scale using transfer learning, which can provide some meaningful information for the researchers who are interested in this field.

We cited Kavzoglu paper in the Introduction:

“Machine learning is currently the most commonly applied method in research for solving the problem of landslide prediction (Goetz et al., 2015; Merghadi et al., 2020; Kavzoglu et al., 2019).”

RC5. The necessity & novelty of the manuscript should be presented and stressed in the “Introduction” section. Please check and refer these papers:

1. Nanehkaran, Y. A., Mao, Y., Azarafza, M., Kockar, M. K., & Zhu, H. H. (2021). Fuzzy-based multiple decision method for landslide susceptibility and hazard assessment: A case study of Tabriz, Iran. *Geomechanics and Engineering*, 24(5), 407-418.

2. Ngo, P. T. T., Panahi, M., Khosravi, K., Ghorbanzadeh, O., Kariminejad, N., Cerda, A., & Lee, S. (2021). Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran. *Geoscience Frontiers*, 12(2), 505-519.

AR5: Thanks for providing the papers to refer.

In the Introduction section, we cited Lin et al. (2021) to clarify the difficulty of establishing the landslide inventory data for model training and testing. Then, we used Sequeira et al. (2016), Wenger and Olden (2012), and Rudy et al. (2016) to verify the model transferability have the potential to help solve the above problem in landslide assessments. According to Yates et al. (2018) study, direct transfer of traditional machine learning models is problematic in landslide studies. A successful model transfer does not necessarily rely solely on the extent of geographic or temporal separation, but rather on the similarity of the environmental conditions between the source and target areas. Meanwhile, following Shimodaira (2000), Pan and Yang (2010), and Yates et al. (2018), model performance can be degraded due to differences in feature space and/or data distributions. Hence, our study aims at solving these problems by using case-based reason and domain adaptation. These references could underline the necessity and novelty of this work in the Intro section.

Because Nanehkaran et al. (2021) and Ngo et al. (2021) studies are not related to transfer learning, we might not cite them in our manuscript. But these great works could be useful for our further study.

References:

1. Lin, Q. G., Lima, P., Steger, S., Glade, T., Jiang, T., Zhang, J. H., Liu, T. X., and Wang, Y.: National-scale data-driven rainfall induced landslide susceptibility mapping for China by accounting for incomplete landslide data, *Geoscience Frontiers*, 12, 101248, <https://doi.org/10.1016/j.gsf.2021.101248>, 2021.

2. Sequeira, A. M. M., Mellin, C., Lozano-Montes, H. M., Vanderklift, M. A., Babcock, R. C., Haywood, M. D. E., Meeuwig, J. J., and Caley, M. J.: Transferability of predictive models of coral reef fish species richness, *Journal of Applied Ecology*, 53, 64-72, <https://doi.org/10.1111/1365-2664.12578>, 2016.

3. Wenger, S. J. and Olden, J. D.: Assessing transferability of ecological models: an underappreciated aspect of statistical validation, *Methods in Ecology and Evolution*, 3, 260-267, <https://doi.org/10.1111/j.2041-210X.2011.00170.x>, 2012.
4. Rudy, A. C. A., Lamoureux, S. F., Treitz, P., and van Ewijk, K. Y.: Transferability of regional permafrost disturbance susceptibility modelling using generalized linear and generalized additive models, *Geomorphology*, 264, 95-108, <https://doi.org/10.1016/j.geomorph.2016.04.011>, 2016.
5. Shimodaira, H.: Improving predictive inference under covariate shift by weighting the log-likelihood function, *J Stat Plan Infer*, 90, 227-244, [https://doi.org/10.1016/S0378-3758\(00\)00115-4](https://doi.org/10.1016/S0378-3758(00)00115-4), 2000.
6. Pan, S. J. and Yang, Q. A.: A Survey on Transfer Learning, *IEEE Transactions on knowledge and data engineering*, 22, 1345-1359, <https://doi.org/10.1109/TKDE.2009.191>, 2010.
7. Yates, K. L., Bouchet, P. J., Caley, M. J., Mengersen, K., Randin, C. F., Parnell, S., Fielding, A. H., Bamford, A. J., Ban, S., Barbosa, A., Dormann, C. F., Elith, J., Embling, C. B., Ervin, G. N., Fisher, R., Gould, S., Graf, R. F., Gregr, E. J., Halpin, P. N., Heikkinen, R. K., Heinanen, S., Jones, A. R., Krishnakumar, P. K., Lauria, V., Lozano-Montes, H., Mannocci, L., Mellin, C., Mesgaran, M. B., Moreno-Amat, E., Mormede, S., Novaczek, E., Opper, S., Crespo, G. O., Peterson, A. T., Rapacciuolo, G., Roberts, J. J., Ross, R. E., Scales, K. L., Schoeman, D., Snelgrove, P., Sundblad, G., Thuiller, W., Torres, L. G., Verbruggen, H., Wang, L., Wenger, S., Whittingham, M. J., Zharikov, Y., Zurell, D., and Sequeira, A. M. M.: Outstanding Challenges in the Transferability of Ecological Models, *Trends in ecology & evolution*, 33, 790-802, <https://doi.org/10.1016/j.tree.2018.08.001>, 2018.

RC6. Provide a literature of the methods used learning-based LSA methods in “Introduction”. The use of a table to demonstrate the advantage disadvantage of these methods can be useful. Towards the end, mention the superiority & repeat the novelty of your work. Please check and refer these papers:

1. Wang, Y., Fang, Z., & Hong, H. (2019). Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China. *Science of the total environment*, 666, 975-993.
2. Fang, Z., Wang, Y., Peng, L., & Hong, H. (2020). Integration of convolutional neural network and conventional machine learning classifiers for landslide susceptibility mapping. *Computers & Geosciences*, 139, 104470.

AR6: Because comparison of modeling techniques is not our objective, dedicating too much space to the presentation of various machine-learning techniques would distract the reader from the main topic of this paper. We used several citations to clarify the popularity of machine learning in landslide assessments and some papers have researched performance of different

methods in landslide assessments, e.g., Goetz et. al (2015) in *Computers & Geosciences*. Hence, readers could find more detailed in those specific papers which are to demonstrate the advantage and disadvantage of learning-based LSA methods.

Reference:

1. Goetz, J. N., Brenning, A., Petschko, H., and Leopold, P.: Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling, *Computers & Geosciences*, 81, 1-11, <https://doi.org/10.1016/j.cageo.2015.04.007>, 2015.

RC7. A relevant source of subjectivity and uncertainty is introduced when splitting the input parameters into an arbitrary number of classes with random break values. These choices affect the results. Would you please describe your solution?

AR7: Thanks for the question. We're not splitting the predictor variables in this study and our results are therefore not affected by such decisions. Instead, we use the GAM, which optimizes the complexity of its transformation functions in a data-driven way using the training data.

RC8. The methodology section is weakly written. So, my suggestion is to reconstruct it. In addition, please kindly use the flowcharts, model verification, benchmarks and error tables to provide detailed methodology.

AR8: Thanks for this comment. We insert a short paragraph to section 2, before subsection 2.1 to further clarify the methodology used in our study:

“In this section we first introduce the general CBR and DA methods separately (sections 2.1 and 2.2). We then explain how CBR and DA models as well as the combined CBR-DA approach are trained and tested in this work (section 2.3). The data used for demonstrating the proposed approaches is then briefly presented, referring the readers to the relevant local literature for details (section 2.4).”

Meanwhile, the flowchart (Figure 1) listed the completed and detailed processes of all methods used in our study. Hence, with the additional statements and flowchart, methodology shows the details better.

RC9. What will be happened if you use this algorithm for another region?

AR9: In our manuscript, we use other regions similar to the target region to obtain a landslide model, and then use this model to make predictions for the target region. We cannot predict with certainty when our algorithm is applied to other regions, but our assessment strategy shows the range of performances we may obtain using a wide range of geographically diverse source and target regions.

For example, in our manuscript, the AUROCs of a model trained by the most related source area Piacenza 25 was 0.762 and that of the model trained with Bologna 25 data itself was also

0.762. Otherwise, on average AUROCs obtained by using our algorithm were below the overoptimistic situation (training and testing in target area) by only  $\sim 0.05$ .

Hence, we might be able to obtain the good result when we use this algorithm for another region.

RC10. Please add a subsection clearly articulating the main limitations, wider applicability of your methods, and findings in the “Discussion” section. Also, the authors should deepen the discussion. Please check and refer these papers:

1. Wu, Y., Ke, Y., Chen, Z., Liang, S., Zhao, H., & Hong, H. (2020). Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping. *Catena*, 187, 104396.
2. Azarafza, M., Ghazifard, A., Akgün, H., & Asghari-Kaljahi, E. (2018). Landslide susceptibility assessment of South Pars Special Zone, southwest Iran. *Environmental Earth Sciences*, 77(24), 1-29.

AR10: Thanks for the comment. We changed the title of section 4.3 to "Utility of domain adaptation in geospatial learning and other limitations". In Discussion, we pointed out limitations for our methods (e.g., "DA did not generally improve transfer learning performance"). It might be clear to describe the limitation for each method used in our paper. For general limitations for all methods used in our study, we added another paragraph in section 4.3:

“Although the study areas cover a wide range of climates with different land cover types and landslide process types, our set of source areas is by no means complete and the results may therefore not be fully representative for the performances that might be achieved at a global scale. Future work should therefore broaden the database of source areas.”

We also thank the reviewer for pointing out the potential of other ML techniques for landslide susceptibility modelling by providing the papers. There are some papers to assess and compare different ML techniques in landslide assessments, for example, Goetz et. al (2015) in *Computers & Geosciences* showed that there's usually not much of a performance difference.

RC11. Please define susceptibility as this terminology is poorly understood by many researchers.

AR11: Thanks for the comment. We added the sentences in Intro section:

“Landslide susceptibility refers to the likelihood of landslides occurring in the region, depending on the local topographic conditions, and estimating the likely location of landslides (Reichenbach et al., 2018).”

RC12. There are no LSA maps that indicate the degree of the susceptibility, please add the required map from triggering factors to risk map. In LSA analysis, providing maps is very important to understand the authors' results.

AR12: The figure 9 showed an example of classified landslide susceptibility maps. Meanwhile, the objective of this study was to evaluate a methodology, not to create susceptibility maps for areas in which these maps already exist from previous studies. Also note that this would require us to include a huge number of maps for a total of 17 target areas and for different modelling strategies, which would not be possible even in supplementary material.

RC13. Please provide justifications for your model, I believe application of other learning based method can be providing much better results (for example: deep learning). Please verify your model with well-known methods.

AR13: Thanks for the suggestion. According to Goetz's publications, we can find GAM also can obtain good results in terms of predictive performance.

Reference:

1. Goetz, J. N., Guthrie, R. H., and Brenning, A.: Integrating physical and empirical landslide susceptibility models using generalized additive models, *Geomorphology*, 129, 376-386, <https://doi.org/10.1016/j.geomorph.2011.03.001>, 2011.

2. Goetz, J. N., Brenning, A., Petschko, H., and Leopold, P.: Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling, *Computers & Geosciences*, 81, 1-11, <https://doi.org/10.1016/j.cageo.2015.04.007>, 2015.

RC14. What type landslide was analyzed? Needs to describe.

AR14: Please find the relevant information in table 3 and we are sorry for not providing more details here, because the cited local literature for each source area could provide more information.

RC15. Please add the “case study” section and provide relevant information about the geology, morphology, seismicity, hydrology and other triggering factors that impact on LSA.

AR15: Please find the section 2.4 Case study transfer source and target areas and table 3. We apologize for not being able to present detailed information in this kind of publication, which can be found in the cited local literature for each source area.

RC16. Please provide the table of hyper-parameters values of all algorithms. Please check and refer these papers:

1. Fang, Z., Wang, Y., Peng, L., & Hong, H. (2021). A comparative study of heterogeneous ensemble-learning techniques for landslide susceptibility mapping. *International Journal of Geographical Information Science*, 35(2), 321-347.

2. Azarafza, M., Azarafza, M., Akgün, H., Atkinson, P. M., & Derakhshani, R. (2021). Deep learning-based landslide susceptibility mapping. *Scientific reports*, 11(1), 1-16.

AR16: Thanks for providing the papers.



A generalized additive model is a generalized linear model with a linear predictor involving a sum of smooth functions of covariates. According to the Wood's book, GAM has no hyperparameters since the complexity of transformation functions is optimized internally (on the training set) in a data-driven way.

Reference:

1. Wood, S. N. (2006). Generalized additive models: an introduction with R. Chapman and Hall/CRC.

RC17. I noticed that the conclusion section tends to repeat abstract and results. The conclusion paragraph should be short, impactful, and direct the reader to this research's next steps and opportunities.

AR17: Thanks for the comment.

For our abstract, the first sentence is to introduce our objective. The rest of it summarizes results and their broader implication, which we believe is appropriate for a Conclusion. Therefore, we describe more details in Conclusion to show reader more clearly about our study's aim, results, and further applications.