nature portfolio

Peer Review File

Machine learning and hydrodynamic proxies for enhanced rapid tsunami vulnerability assessment

Web links to the author's journal account have been redacted from the decision letters as indicated to maintain confidentiality

1st Feb 24

Dear Dr Scorzini,

Please allow us to apologise for the delay in sending a decision on your manuscript titled "Machine learning and hydrodynamic proxies for enhanced rapid tsunami vulnerability assessment". It has now been seen by 3 reviewers, and we include their comments at the end of this message. They find your work of interest, but some important points are raised. We are interested in the possibility of publishing your study in Communications Earth & Environment, but would like to consider your responses to these concerns and assess a revised manuscript before we make a final decision on publication.

We therefore invite you to revise and resubmit your manuscript, along with a point-by-point response that takes into account the points raised. Please highlight all changes in the manuscript text file.

We are committed to providing a fair and constructive peer-review process. Please don't hesitate to contact us if you wish to discuss the revision in more detail.

Please use the following link to submit your revised manuscript, point-by-point response to the referees' comments (which should be in a separate document to any cover letter), a tracked-changes version of the manuscript (as a PDF file) and the completed checklist:
[link redacted]
** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first **

We hope to receive your revised paper within six weeks; please let us know if you aren't able to submit it within this time so that we can discuss how best to proceed. If we don't hear from you, and the revision process takes significantly longer, we may close your file. In this event, we will still be happy to reconsider your paper at a later date, as long as nothing similar has been accepted for publication at Communications Earth & Environment or published elsewhere in the meantime.

Please do not hesitate to contact us if you have any questions or would like to discuss these revisions further. We look forward to seeing the revised manuscript and thank you for the opportunity to review your work.

Best regards,

Olusegun Dada, PhD
Editorial Board Member
Communications Earth & Environment
orcid.org/0000-0001-8020-3323

Joe Aslin
Deputy Editor
Communications Earth & Environment

EDITORIAL POLICIES AND FORMATTING

We ask that you ensure your manuscript complies with our editorial policies. Please ensure that the following formatting requirements are met, and any checklist relevant to your research is completed and uploaded as a Related Manuscript file type with the revised article.

Editorial Policy: Policy requirements (Download the link to your computer as a PDF.)

Furthermore, please align your manuscript with our format requirements, which are summarized on the following checklist:
Communications Earth & Environment formatting checklist

and also in our style and formatting guide Communications Earth & Environment formatting guide .

*** DATA: Communications Earth & Environment endorses the principles of the Enabling FAIR data project (http://www.copdess.org/enabling-fair-data-project/). We ask authors to make the data that support their conclusions available in permanent, publically accessible data repositories. (Please contact the editor if you are unable to make your data available).

All Communications Earth & Environment manuscripts must include a section titled "Data Availability" at the end of the Methods section or main text (if no Methods). More information on this policy, is available at http://www.nature.com/authors/policies/data/data-availability-statements-data-citations.pdf.

In particular, the Data availability statement should include:
- Unique identifiers (such as DOIs and hyperlinks for datasets in public repositories)
- Accession codes where appropriate
- If applicable, a statement regarding data available with restrictions
- If a dataset has a Digital Object Identifier (DOI) as its unique identifier, we strongly encourage including this in the Reference list and citing the dataset in the Data Availability Statement.

DATA SOURCES: All new data associated with the paper should be placed in a persistent repository where they can be freely and enduringly accessed. We recommend submitting the data to discipline-specific, community-recognized repositories, where possible and a list of recommended repositories is provided at http://www.nature.com/sdata/policies/repositories.

If a community resource is unavailable, data can be submitted to generalist repositories such as figshare or Dryad Digital Repository. Please provide a unique identifier for the data (for example a DOI or a permanent URL) in the data availability statement, if possible. If the repository does not provide identifiers, we encourage authors to supply the search terms that will return the data. For data that have been obtained from publically available sources, please provide a URL and the specific data product name in the data availability statement. Data with a DOI should be further cited in the methods reference section.

Please refer to our data policies at http://www.nature.com/authors/policies/availability.html.

REVIEWER COMMENTS:

Reviewer #1 (Remarks to the Author):

The study discusses the importance of additional variables for tsunami fragility curves, particularly those related to the flow velocity. Unlike the flow depth or inundation height, data on flow velocity is limited, if any. Thus, most of the existing studies relied on the simulation result. However, the lack of flow velocity data makes evaluating the model's performance difficult. Therefore, proxies to this variable are necessary to complement the simulation. The study also highlights the interpretability of their machine learning model, which is one of the appealing factors of this paper. The manuscript is well written, i.e., the objective and overall description are clear and concise. Therefore, I would recommend this paper for publication after addressing the following comments.

1. It is evident from Figure 1 that the contribution of simulated flow velocity in most of the models is minimal relative to its proxies. Is the simulated flow velocity from the backwash incorporated, or can it be distinguished in this analysis? The force from tsunami backwash can sometimes be as strong or even more potent than the initial impact.

2. Page 3 Line 2. De Risi et al. (2017) [11] study does not use the stated empirical relation to estimate the flow velocity. This citation also appears on Page 13, Line 11. Please check again (I might be wrong though).

3. Page 6 Line 1. 0.820 → 0.828.

4. Page 6 Line 27. wave velocity → flow velocity. Wave velocity can be misunderstood for phase velocity, which differs from current (flow velocity).

5. The density plots of Figure 2 are not discussed in the text (only correlation analysis between variables is mentioned). A brief description should suffice to explain the meaning of these plots.

6. It will be useful to plot the data overlaid with the resulting fragility curve, at least for some examples, to see the performance or goodness-of-fit of the models.

7. Please define RC in Figure 5.

8. Page 14 Line 20 and Figure 6. The paper refers to Di Bacco et al. (2023) for defining buffer geometry. However, it is better to state the definition again here. Why does the buffer zone have a particular shape, like that plotted in Figure 6? What is the gray-colored area at the bottom left corner of the plots? Is it the sea?

9. Page 16 Line 16. Please explain why sediment transport modeling is required in the tsunami simulation—no variable listed in Table 1 is related to the property of sediment.

10. Page 17 Line 24. What justifies that 100 ensemble predictions are sufficient to represent the uncertainty level? Is there a way to quantify such a hyperparameter?

Reviewer #2 (Remarks to the Author):

The manuscript by Scorzini et al. investigates the interplay of factors impacting building vulnerability to tsunamis, with a specific emphasis on the hydrodynamic effects linked to tsunami propagation on land. The study comprehensively explores various building- and site-related features that, despite their recognized significance in the literature, have been largely disregarded in current damage models.

(1) Regarding the results of synthetic variables for shielding mechanism and debris impact as proxies for water velocity, the lack of apparent correlation between vsim and NShArea, as mentioned in the text, may be contradictory to intuitive understanding in physics. Despite the existence of the shielding effect arising from the presence of buildings, are there other factors that may indirectly influence it?
(2) Meanwhile, the comparison between models ID2 and ID6 suggests that the addition of vc (velocity variable) does not lead to performance improvement. This may indicate the presence of redundancy among variables in the model, where certain variables provide overlapping information, rendering their contributions to the model without actual gain. Such redundancy can make the model more challenging to interpret and optimize. I suggest that authors should discuss how to address such redundancy during the model comparison.
(3) The study raises concerns about the black-box nature of machine learning algorithms. How can model interpretability be enhanced for a better understanding of the predictions, especially in practical applications of tsunami damage prediction?
(4) I agree that Figure 4 well represents the amplification effect induced by ria-type coasts, yet does it consider additional geographical differences, such as variations in geological conditions? Are these differences adequately accounted for in the model? At the same time, I also believe that a more quantitative approach is needed for the comparison of different morphological variables.

In addition, I have a few minor comments:
1) Page 2, Line 12: I suggest that the authors could provide a specific description of the manifestation of fragility functions.
2) Page 2, Line 23: Estimating tsunami velocity is challenging. In the future, is it possible to observe tsunami velocity by combining it with HF radar?
3) Page 3, Line 11: Here you can briefly review the application of deep learning in tsunami prediction. The following two articles can be referred to.
https://doi.org/10.1186/s40623-023-01912-6
https://doi.org/10.1038/s41467-022-33253-5
4) Page 14, Line 11: What are the conditions for the application of this empirical formula, and what are its limitations? Please provide a clear explanation.
5) Page 16, Line 7: Change 97’852 to 97,852.
6) Page 16, Line 23: How come the sum of these three parts is not equal to 100%? Or does it mean that the validation set is also used as the test set?

Reviewer #3 (Remarks to the Author):

Overall the paper is interesting; however, major revisions are required before publication. See the following comments and feedback for your consideration:

- Page 3: Additional background literature is needed for the application of machine learning methods in damage assessments. The authors are primarily emphasizing the findings of Di Bacco et al.'s study (2023).
- Page 4: What does MLIT stand for?
- Page 5: It is not clear to me how "essential variables" are determined. Clarification would be helpful.
- Figure 1: Accuracy score is selected as a metric to evaluate predictive performance. Did you encounter an imbalanced dataset issue in your case? The F1-score might be a more appropriate measure to consider if this is the case. It would be important to elaborate on this matter in the manuscript.
- Figure 1: There is a brief discussion about the correlation among some of the input variables. However, It would be insightful if the authors could elaborate on this and discuss it quantitively, as it can influence the conclusions drawn here. For instance, what is the relative importance of input variables in different combination scenarios? Does their sum add up to one?
- What are the limitations of your methodology and the barriers toward the widespread adoption of multi-variate models in tsunami vulnerability assessment? It would be useful to discuss this in the conclusions.
- Methods: It would be useful to clearly indicate the measures you took to prevent bias and overfitting in your machine learning models. Have you performed hyperparameter tuning?

We would like to thank the Reviewers for their interest in our work and for carefully reading our manuscript; we greatly appreciate their insightful comments as they contribute to increase the manuscript robustness and, in general, to improve its quality. We provide below a point-by-point reply to the general and specific comments raised.
You can see all the changes made on the manuscript in the file “Scorzini_et_al_rev_track_changes”. The “clear” version of the revised manuscript is in the file “Scorzini_et_al_rev”.

REVIEWER 1

The study discusses the importance of additional variables for tsunami fragility curves, particularly those related to the flow velocity. Unlike the flow depth or inundation height, data on flow velocity is limited, if any. Thus, most of the existing studies relied on the simulation result. However, the lack of flow velocity data makes evaluating the model’s performance difficult. Therefore, proxies to this variable are necessary to complement the simulation. The study also highlights the interpretability of their machine learning model, which is one of the appealing factors of this paper. The manuscript is well written, i.e., the objective and overall description are clear and concise. Therefore, I would recommend this paper for publication after addressing the following comments.

R1.C1. It is evident from Figure 1 that the contribution of simulated flow velocity in most of the models is minimal relative to its proxies. Is the simulated flow velocity from the backwash incorporated, or can it be distinguished in this analysis? The force from tsunami backwash can sometimes be as strong or even more potent than the initial impact.

Reply: The hydrodynamic simulation accounted for the entire duration of the event, including both the main wave and the backwash. As stated in the original manuscript (P16.L9-12), at each building location, we considered the maximum water depth and flow velocity values calculated over the simulation. These values inherently include both the effects, even though they are not explicitly distinguished in our analysis. To enhance clarity, we have enhanced the description of the hydrodynamic data as follows (P17.L19-24 in the revised manuscript): “The simulations reproduced tsunami propagation and inundation for 3 hours from the earthquake, then capturing the effects of both the primary wave and the subsequent backwash. Inundation depth (h_{sim}) and velocity (v_{sim}) values sampled at the buildings’ location represent the maximum estimated values from the simulation; this approach is based on the assumption that observed damage is triggered by overall maximum demands, despite the potential variation in the timing of occurrence for the maximum depth and velocity at a given location.”

R1.C2. Page 3 Line 2. De Risi et al. (2017) [11] study does not use the stated empirical relation to estimate the flow velocity. This citation also appears on Page 13, Line 11. Please check again (I might be wrong though).

Reply: Thank you for your observation. Upon further review, we acknowledge that there was an error in the citation on P14.L11. The correct citation should indeed be [15] and we then have rectified this in the revised version of the manuscript. However, we would like to clarify that the reference to De Risi et al. (2017) [11] on P3.L2, was intentional. In this context, we referred to their work within a broader literature context, as they reported the theoretical relation for shallow water between inundation depth and flow velocity based on observations from the 2011 Tohoku tsunami. We have retained this reference to provide a comprehensive overview of the relevant literature.

R1.C3. Page 6 Line 1. 0.820 → 0.828.

Reply: Fixed.

R1.C4. Page 6 Line 27. wave velocity → flow velocity. Wave velocity can be misunderstood for phase velocity, which differs from current (flow velocity).

Reply: Thank you for pointing out the potential confusion. We have revised the term from “wave velocity” to “flow velocity” to ensure clarity and consistency throughout the manuscript.

R1.C5. The density plots of Figure 2 are not discussed in the text (only correlation analysis between variables is mentioned). A brief description should suffice to explain the meaning of these plots.

Reply: In the revised version of the manuscript (P6.L1-5), we have included a description of the features and aim of the pair plot shown in Figure 2. In addition, we have also leveraged this plot for discussing the possible effects of the imbalance between the damage classes on the results of the study (see also reply to R3.C4).

R1.C6. It will be useful to plot the data overlaid with the resulting fragility curve, at least for some examples, to see the performance or goodness-of-fit of the models.

Reply: In our approach, as already described in the original methodological section, fragility functions are directly derived from the output of the XT model and are not obtained by fitting a functional form to data points, as in traditional empirical functions. In this sense, the models' accuracy (Figure 1, as well as in the confusion matrices included in the new Supplement) already provides a comprehensive information on model's uncertainty.

R1.C7. Please define RC in Figure 5.

Reply: In the revised version of the manuscript we have included the definition of "RC" (i.e., Reinforced Concrete) where it is first mentioned in the text (P13.L4 of the revised manuscript).

R1.C8. Page 14 Line 20 and Figure 6. The paper refers to Di Bacco et al. (2023) for defining buffer geometry. However, it is better to state the definition again here. Why does the buffer zone have a particular shape, like that plotted in Figure 6? What is the gray-colored area at the bottom left corner of the plots? Is it the sea?

Reply: Following Reviewer's suggestion, in the revised version of the methodological section, we have included details on the criterion established by Di Bacco et al. (2023) for defining buffer geometry. In detail, the following text has been included (P16.L4-10): "Di Bacco et al. [23] established a criterion for the definition of the buffer geometry drawing inspiration from the work of Naito et al. [27], who assumed debris propagation from dispersed vessels toward a direction perpendicular to the coast, with a 45° spread angle that could increase in areas near the coast due to drawdown. In detail, these polygons are characterized by extreme points represented by the centroid of the examined building and the nearest point on the coastline. Di Bacco et al. tested two different shapes, finding no substantial effects of buffer geometry on the accuracy of developed damage models. In the present study, we then opted for their proposed narrower geometry, as exemplified in Figure 6."
In addition, we have revised Figure 6 by including a label that identifies the coastline (and then the sea) to make it more clear.

R1.C9. Page 16 Line 16. Please explain why sediment transport modeling is required in the tsunami simulation - no variable listed in Table 1 is related to the property of sediment.

Reply: Simulation results of the event utilized in this study included sediment transport phenomena, a key component for an accurate representation of tsunami hydrodynamics (i.e., water depth and flow velocity), as already underlined in the original manuscript (P15.L11-12). However, we opted not to incorporate any sediment-related variables in the development of the damage models due to the infrequent availability of such data in common tsunami simulations (due to required input data, computational times and overall "modelling costs"), which makes it practically unfeasible to integrate into a damage prediction framework.

R1.C10. Page 17 Line 24. What justifies that 100 ensemble predictions are sufficient to represent the uncertainty level? Is there a way to quantify such a hyperparameter?

Reply: The 100 iterations for obtaining the fragility functions with the corresponding uncertainty bounds were defined in a preliminary assessment of the stability of the results, while keeping feasible computational times.

REVIEWER 2

The manuscript by Scorzini et al. investigates the interplay of factors impacting building vulnerability to tsunamis, with a specific emphasis on the hydrodynamic effects linked to tsunami propagation on land. The study comprehensively explores various building- and site-related features that, despite their recognized significance in the literature, have been largely disregarded in current damage models.

R2.C1. Regarding the results of synthetic variables for shielding mechanism and debris impact as proxies for water velocity, the lack of apparent correlation between vsim and NShArea, as mentioned in the text, may be contradictory to intuitive understanding in physics. Despite the existence of the shielding effect arising from the presence of buildings, are there other factors that may indirectly influence it?

Reply: In reply to the observation made by the Reviewer, it is important to clarify the distinction between numerical correlation and a direct cause-and-effect relationship. The absence of a strong numerical correlation between v_{sim} and *NShArea* does not negate the existence of the shielding effect due to buildings. As already stated in the original manuscript (P6.L21-25), buildings can be seen as macro-scale roughness elements, introducing changes in bottom friction from the hydrodynamic model perspective. It is this alteration in bottom friction that influences the flow patterns of tsunami propagation on land, reinforcing our claim that shielding serves as a proxy for flow velocity. This clarification has been further articulated in the revised manuscript (P7.L7-9).
The Reviewer is right in pointing out that various factors (such as the distance from the coast, inundation depth, and others) play a role in influencing flow velocity. However, we would also like to emphasize that, while it is possible to observe a direct relationship between the variables for an individual building under different inundation events and shielding configurations, the same direct relationship may not be readily apparent for a single event on a large spatial scale, as in the examined case.

R2.C2. Meanwhile, the comparison between models ID2 and ID6 suggests that the addition of vc (velocity variable) does not lead to performance improvement. This may indicate the presence of redundancy among variables in the model, where certain variables provide overlapping information, rendering their contributions to the model without actual gain. Such redundancy can make the model more challenging to interpret and optimize. I suggest that authors should discuss how to address such redundancy during the model comparison.

Reply: We would like to highlight that the issue of redundancy has been already addressed in the original manuscript (P6.L1-5). Specifically, the variable v_c, being a direct function of h, does not contribute additional information to the model and, therefore, such variable can be simply ignored in model development, since the same information is already provided by h.

R2.C3. The study raises concerns about the black-box nature of machine learning algorithms. How can model interpretability be enhanced for a better understanding of the predictions, especially in practical applications of tsunami damage prediction?

Reply: Our approach centers around "explanation through visualization", aiming to provide explicit and insightful interpretations in the face of the inherent complexity associated with multivariable problems in tsunami damage prediction. As discussed in the paper, our approach is dedicated to offering a more informative understanding of the damage mechanisms, despite the apparent "black-box" nature of machine-learning algorithms. To enhance model interpretability, we have chosen to translate the results of our models into traditional fragility functions. This not only facilitates a physical explanation of the role of different features on damage, but also allows us to emphasize the limitations of uni-variable damage models through the visualization of uncertainty bounds.

R2.C4. I agree that Figure 4 well represents the amplification effect induced by ria-type coasts, yet does it consider additional geographical differences, such as variations in geological conditions? Are these differences adequately accounted for in the model? At the same time, I also believe that a more quantitative approach is needed for the comparison of different morphological variables.

Reply: We appreciate Reviewer's suggestion regarding the inclusion of additional geographical differences, such as variations in geological conditions. However, we would like to note that our model still implicitly accounts for macro-geological variations through the inclusion of the variable *CoastType*, which distinguishes between ria and plain coasts, aligning with the broader geological distinctions in the Tohoku region (as in the geological map of Japan, reported, for instance, in Hartmann et al. (2010, doi: 10.1007/s00531-008-0381-5)). In this sense, introducing an additional macro-geological variable would not necessarily lead to improved performance (due to the overlap with *CoastType*), similarly to the observed case for the variable v_c (as discussed in response to R2.C2). Moreover, we would like to mention that our research group is instead currently engaged in another study, which will contribute to a future publication, aiming at evaluating more micro-scale effects of soil characteristics on tsunami-induced damage to various assets.

Minor comments:

R2.C5. Page 2, Line 12: I suggest that the authors could provide a specific description of the manifestation of fragility functions.

Reply: Upon reviewing the text, we believe that a description of fragility functions and their types has been already provided in the original manuscript (P2.L8-18). Additionally, we cited reference papers that discuss tsunami fragility functions. If there are specific aspects the Reviewer would like us to elaborate on or if there are particular details he/she finds missing, please provide further guidance so that we can address such concerns more effectively.

R2.C6. Page 2, Line 23: Estimating tsunami velocity is challenging. In the future, is it possible to observe tsunami velocity by combining it with HF radar?

Reply: Despite not being experts in HF radar technology, we acknowledge its potential for the estimation of tsunami velocity. However, we think that putting emphasis on this topic might be beyond the scope of our paper, which is focused on understanding the effects of various factors on damage mechanisms (with a focus on velocity-related ones), rather than discussing the detailed methods for velocity measurement. Based on this consideration, we do not plan to add comments on this aspect to maintain the alignment with the main aim of our paper.

R2.C7. Page 3, Line 11: Here you can briefly review the application of deep learning in tsunami prediction. The following two articles can be referred to.

https://doi.org/10.1186/s40623-023-01912-6

https://doi.org/10.1038/s41467-022-33253-5

Reply: Thank you for your valuable suggestion. Upon careful consideration of the two papers, we observed that they focus on tsunami inundation prediction, a topic beyond the scope of our study and the main message of the introduction, dealing with the multi-variable nature of tsunami damage. In this context, we have therefore opted not to include either of the suggested articles.

R2.C8. Page 14, Line 11: What are the conditions for the application of this empirical formula, and what are its limitations? Please provide a clear explanation.

Reply: As stated in the original manuscript and supported by cited references, the theoretical basis for this formula is rooted in Bernoulli's theorem. The only variation among different empirical formulas in the literature lies in the value of the coefficient ($v = c\sqrt{gh}$), with c typically ranging between 0.7 and 2.0. In a machine-learning approach, the specific value of a constant does not have any impact on the results and, therefore, we selected 0.7 as reference value. We have clarified this aspect in the revised version of the manuscript (P15.L12-16).

R2.C9. Page 16, Line 7: Change 97'852 to 97,852.

Reply: Fixed.

R2.C10. Page 16, Line 23: How come the sum of these three parts is not equal to 100%? Or does it mean that the validation set is also used as the test set?

Reply: We suspect that the Reviewer might have misunderstood the description, which we think is already quite explicative. We here clarify for the Reviewer that 5% is the proportion of the validation dataset; the remaining 95% is then further divided into two part: 95% (i.e., 90.25% of the total) is the training set and 5% (i.e., 4.75% of the total) is the test set.

REVIEWER 3

Overall the paper is interesting; however, major revisions are required before publication. See the following comments and feedback for your consideration:

R3.C1. Page 3: Additional background literature is needed for the application of machine learning methods in damage assessments. The authors are primarily emphasizing the findings of Di Bacco et al.'s study (2023).

Reply: Our primary focus has been on referencing machine learning studies specifically related to tsunami damage and Di Bacco et al.' s work is a significant contribution in this context, also considering the fact that it introduced the proxy variables for shielding and debris impact analyzed in detail in this study. Acknowledging the broader applicability of machine learning for damage modelling also for other natural hazards, we recognize the importance of a more comprehensive literature review. In response to Reviewer's suggestion, we have then included in the revised manuscript references to relevant studies on machine learning applied to earthquake, hurricane and flood damage modeling, including:

Saengtabtim, K., Leelawat, N., Tang, J., Treeranurat, W., Wisittiwong, N., Suppasri, A., Pakoksung, K., Imamura, F., Takahashi, N., Charvet, I. Predictive analysis of the building damage from the 2011 Great East Japan tsunami using decision tree classification related algorithms. IEEE Access 9 31065-31077 (2021).

Xie, Y., Ebad Sichani, M., Padgett, J.E. & DesRoches, R. The promise of implementing machine learning in earthquake engineering: A state-of-the-art review. Earthq. Spectra 36(4), 1769-1801 (2020).

Szczyrba, L., Zhang, Y., Pamukcu, D., Eroglu, D. I., & Weiss, R. Quantifying the role of vulnerability in hurricane damage via a machine learning case study. Nat. Hazards Rev. 22(3), 04021028 (2021).

Alipour, A., Ahmadalipour, A., Abbaszadeh, P., & Moradkhani, H. Leveraging machine learning for predicting flash flood damage in the Southeast US. Environ. Res. Lett 15(2), 024011 (2020).

R3.C2. Page 4: What does MLIT stand for?

Reply: MLIT stands for Ministry of Land, Infrastructure, Transport and Tourism of Japan. To ensure clarity, we have explicitly defined this acronym the first time it appears in the revised version of the manuscript (P3.L13).

R3.C3. Page 5: It is not clear to me how "essential variables" are determined. Clarification would be helpful.

Reply: Thank you for the opportunity for clarification. With "essential variables" we meant "basic" or "minimum information" variables available in the context of our study, which consisted in the features reported in the original MLIT dataset and simple geospatial variables proposed by Di Bacco et al (2023), such as *Distance* and *WDir*. We have then adjusted the terminology for improved clarity in the revised version of the manuscript, as follows: "This baseline incorporates only basic input variables sourced from the original MLIT database, further enriched with some of the geospatial variables introduced by Di Bacco et al. characterized by the most straightforward computation [23]."

R3.C4. Figure 1: Accuracy score is selected as a metric to evaluate predictive performance. Did you encounter an imbalanced dataset issue in your case? The F1-score might be a more appropriate measure to consider if this is the case. It would be important to elaborate on this matter in the manuscript.

Reply: For more comprehensive insights on the results, we have included a Supplement to the revised version of the paper, reporting the confusion matrices for the different developed models. Relevant text has been incorporated into the methodological and results sections of the revised paper, as appropriate (P7.L20-26 and P18.L10-12).

We think that confusion matrices can be even more informative than the F1-score to depict the error patterns across the different damage classes. The Reviewer can note from the matrices, along with the pair plot in Figure 2, the dataset's imbalance for class 5, which is underrepresented. In line with the primary focus of our paper (comparing the effect of velocity-related variables), we opted not to implement any procedure for imbalance management, since it affects in the same way all models trained with various feature combinations, as also evident from the confusion matrices. This aspect has been also clarified in the revised manuscript (P7.L21-26).

R3.C5. Figure 1: There is a brief discussion about the correlation among some of the input variables. However, It would be insightful if the authors could elaborate on this and discuss it quantitively, as it can influence the conclusions drawn here. For instance, what is the relative importance of input variables in different combination scenarios? Does their sum add up to one?

Reply: First, we think that the pair plot shown in Figure 2, which also distinguishes the different damage classes, provides much more informative insights than a single correlation coefficient value, whose discussion would not be particularly meaningful, due to very low significance of the correlations identified among the different variables. Additionally, it is worth noting that the original paper already discussed the impact of correlation on feature importance; this is evident when related variables are added in the model, resulting in no improvement in the overall accuracy and a decrease in the importance associated with each of them. This has been done in the original paper for those variables reporting a more evident relationship and effect (h and v_c). In the revised paper, we have also included a comment on a second example regarding *NShArea* and *NDIArea* (P6.L18-21), while for other combinations we feel that inclusion of further discussion would not add much to the results, given the very weak relationships highlighted in Figure 2.
Regarding the comment on the relative importance of the features, we would like to clarify that Equation 3 has been used in our study to rank the variables based on their effect on mean decrease in accuracy (mda) when they are randomly shuffled. Relative importance would be expressed as the ratio of each mda to the maximum mda value within the considered features. However, we think that displaying results in terms of relative feature importance would offer less information than absolute mda values. In Figure 1, the sum of mda for each model is not 1 because it directly represents the mda induced by each feature. Nevertheless, the ranking of the features or their "relative importance" can be still inferred by comparing the different sizes of the circles, as already mentioned in the original paper (P5.L11-13: "Insights into the importance of individual input features on the models' predictive performance are provided by the circles, the size of which corresponds to the mean decrease in accuracy (mda) when each single variable is randomly shuffled" and legend of Figure 1).

R3.C6. What are the limitations of your methodology and the barriers toward the widespread adoption of multi-variate models in tsunami vulnerability assessment? It would be useful to discuss this in the conclusions.

Reply: We have included a discussion on the main barrier to a widespread adoption of machine learning-based approaches for tsunami damage modelling in the revised version of the conclusions, as follows: "[…] In this context, while the presented modeling framework holds considerable promise for replication across diverse geographical regions and contexts, a primary barrier to its widespread application lies in the scarcity of comprehensive ex-post damage datasets for both model training for new models and validation for a spatial transfer of models developed across diverse regions. This challenge highlights not only the critical need for the establishment of standardized procedures for collecting damage and ancillary data in the aftermath of tsunami events, but also the importance of an open distribution of resulting datasets within the research community." (P14.L26 to P15.L4 of the revised manuscript).

R3.C7. Methods: It would be useful to clearly indicate the measures you took to prevent bias and overfitting in your machine learning models. Have you performed hyperparameter tuning?

Reply: We avoided overfitting in our machine learning models by evaluating accuracy on an independent test set, distinct from the training data (refer to P16.L21-23 in the original manuscript). Additionally, the presented results are averaged over ten training iterations (P5.L9 in the original manuscript) and the extensive dataset size, coupled with the relatively conservative number of features, further safeguards against overfitting. Concerning hyperparameter tuning, a validation set was specifically employed, as detailed in the methodological section (P16.L26-27 in the original manuscript).

11th Apr 24

Dear Dr Scorzini,

Please allow us to apologise for the delay in sending a decision on your manuscript titled "Machine learning and hydrodynamic proxies for enhanced rapid tsunami vulnerability assessment". It has now been seen by our reviewers, whose comments appear below. In light of their advice we are delighted to say that we are happy, in principle, to publish a suitably revised version in Communications Earth & Environment under the open access CC BY license (Creative Commons Attribution v4.0 International License).

We therefore invite you to revise your paper one last time to address the remaining concerns of our reviewers. At the same time we ask that you edit your manuscript to comply with our format requirements and to maximise the accessibility and therefore the impact of your work.

EDITORIAL REQUESTS:

Please review our specific editorial comments and requests regarding your manuscript in the attached "Editorial Requests Table".

*****Please take care to match our formatting and policy requirements. We will check revised manuscript and return manuscripts that do not comply. Such requests will lead to delays. *****

Please outline your response to each request in the right hand column. Please upload the completed table with your manuscript files as a Related Manuscript file.

If you have any questions or concerns about any of our requests, please do not hesitate to contact me.

SUBMISSION INFORMATION:

In order to accept your paper, we require the files listed at the end of the Editorial Requests Table; the list of required files is also available at https://www.nature.com/documents/commsj-file-checklist.pdf .

OPEN ACCESS:

Communications Earth & Environment is a fully open access journal. Articles are made freely accessible on publication under a [CC BY license](http://creativecommons.org/licenses/by/4.0/) (Creative Commons Attribution 4.0 International License). This license allows maximum dissemination and re-use of open access materials and is preferred by many research funding bodies.

For further information about article processing charges, open access funding, and advice and support from Nature Research, please visit https://www.nature.com/commsenv/article-processing-charges

At acceptance, you will be provided with instructions for completing this CC BY license on behalf of all authors. This grants us the necessary permissions to publish your paper. Additionally, you will be asked to declare that all required third party permissions have been obtained, and to provide billing information in order to pay the article-processing charge (APC).

Please use the following link to submit the above items:
[link redacted]
** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first **

We hope to hear from you within two weeks; please let us know if you need more time.

Best regards,

Olusegun Dada
Editorial Board Member
Communications Earth & Environment

Joe Aslin
Deputy Editor,
Communications Earth & Environment
https://www.nature.com/commsenv/
Twitter: @CommsEarth

REVIEWERS' COMMENTS:

Reviewer #1 (Remarks to the Author):

I appreciate the author's effort in revising the manuscript, and my comments have been sufficiently addressed in the revised version. Therefore, I am happy to recommend the publication of this manuscript.

Reviewer #2 (Remarks to the Author):

The manuscript was greatly improved after the authors' revision. I thank them in carefully responding to me and other two reviewers.
While I agree with their most response, I still have questions regarding the author's response in R2.C3. The author claims their method involves explanation through “visualization”, but I haven't found where visualization is specifically manifested in the manuscript. Could the author please provide a detailed explanation?
Thank you!

Reviewer #3 (Remarks to the Author):

I would like to thank the authors for thoroughly addressing my comments. I am satisfied with their responses and have no further comments.

REVIEWER 2

R2.C1. The manuscript was greatly improved after the authors' revision. I thank them in carefully responding to me and other two reviewers. While I agree with their most response, I still have questions regarding the author's response in R2.C3. The author claims their method involves explanation through "visualization", but I haven't found where visualization is specifically manifested in the manuscript. Could the author please provide a detailed explanation?

Reply: We are grateful for acknowledgment of the improvements made to the manuscript. In addressing the final comment, we aim to provide a clear understanding of our approach to "explanation through visualization". Our method originates from the recognition of the inherent challenge posed by the interpretability of machine learning model outputs. Given their lack of direct explainability, we bridged this gap by demonstrating how the output generated by such models can be effectively visualized to convey valuable insights.
To achieve this, we have translated the results of our trained models into the form of traditional fragility functions, as shown in Figures 3, 4, and 5. These functions illustrate the probability of exceeding a certain damage state as a function of inundation depth. Detailed discussions on these visual representations are provided from Page 9, Lines 5-7 onwards, with the methodology for obtaining them elaborated in the "Methods" section from Page 18, Line 20 of the manuscript.
Through the examination of these functions under various conditions, readers can gain a deeper understanding of the relationships between variables, thereby enhancing the interpretability and utility of our methodology.