Reply to Referee 1 (Annotation), Pedro Costa (NHESS)
Major comments:

Interactive comment on "Reconstruction of flow conditions from 2004 Indian Ocean tsunami deposits at the Phra Thong island using a deep neural network inverse model" by Rimali Mitra et al.,

We thank the reviewer for the critical assessment of our manuscript and for the numerous comments and suggestions. Please find our responses to each comment below (in bold italics).

1. The title does not clearly correspond to the content. Either the authors change the manuscript accordingly and provide clear field validation information or they must remove "tsunami deposits" from the title.

   RE: Thank you for your comment. In accordance with your comment, we will add a figure showing the detailed information of the analyzed sediments. We have used the field data from Phra Thong island, Thailand to reconstruct the flow conditions of 2004 Indian Ocean tsunami. We have included our dataset as "Thai_gs5.csv" in https://doi.org/10.5281/zenodo.4075137. The codes include the same dataset to predict the final results. So, we decided to keep the title unchanged.

2. Are you referring to grain-size? When I look at Figure 9 that is what you are presenting. One thing in sediment concentration and grain-size distribution on the incoming tsunami waves, another totally different is grain-size and packing on the tsunami deposits. Throughout the text is not clear what you are determining, because it was not possible how you estimate sediment concentration on the wave from the deposit if your are solely relying on the inverse model. So, this needs to be clarified and the fact that you did not incorporated any reference to "sediment concentration " on the (minimal) conclusions provide further increases my confusion.

   RE: Thank you for the comment. We will add a new diagram to show thickness and grain size distribution from the study area. Here, we refer to the sediment concentration but in Figure 9 we presented volume per unit area and spatial grain-size distribution. We reconstructed sediment concentration of the tsunami deposits.
   In response to your comment we will modify the text in the discussion part and added Goto et al. (2014) reference for clarification in our revised manuscript.
   Our inversion model estimates values of sediment concentration of tsunami that best explain distribution of thickness and grain size of tsunami deposits. The inverse model is trained from the results of the forward model calculation.

3. Please rewrite. How did you determined the post-tsunami concentration? Was this data based on post-tsunami survey? Please make the text more fluent.

   Thank you for the suggestion. We determined the post-tsunami concentration using the DNN inverse model, which automatically finds the distribution of sediment concentration that best

realizes the actual observed distribution of thickness and grain size of the tsunami deposit. We will modify the text in our revised manuscript.
Revised text: The DNN inverse analysis reconstructed the values of flow conditions such as maximum inundation length, flow velocity and maximum flow depth, sediment concentration of five grain size classes using the datasets of grain-size distribution measured from the post-tsunami survey around Phra Thong island.

4. Abstract, Agree but the greater challenge is to study older deposits and reconstruct physical parameters from them.

   RE: We agree with the reviewer.

5. P-1, Line 15, This sentence is somewhat confusing. Please rewrite.

   RE: Thank you for the suggestion. We will modify the text in the revised manuscript.

6. P-2, Line 25, Please add citation.

   RE: Thank you for identifying this. We have added the citation.

7. Line 35, Please revise English. The wroding and figures are not well-structured.

   RE: Thank you for the suggestion. We have revised the text.

8. Please see Costa et al., 2011, Earth Surface Processes and Landforms, Costa et al., 2012, The Holocene and Moreira et al., 2017, Marine Geology

   RE: Thank you for the suggestion. We considered adding costa et al., 2011 and Moreira et al., 2017.

9. P-3, Line 73, The big problem with this manuscript is that is not clear how the inverse model was validated. Figure 9 shows something but much further detail needs to be provided much earlier in the text for the reader to understand what the authors achieved or are aiming.

   RE: Thank you for the comment. We will provide a paragraph on the workflow of the inverse model for detailed clarifications.
   P.3, Line 73, Here, we conduct an DNN inverse analysis of the tsunami deposits measured at Phra Thong island and reconstruct the flow conditions such as the maximum inundation length, flow velocity, maximum flow depth and sediment concentrations of five grain-size classes. The inverse model was based on the forward model, which was proposed by Naruse and Abe, (2017). The forward model calculations were iterated at random initial flow conditions to produce artificial training data sets that represent depositional characteristics such as the spatial distribution of thickness and grain-size composition. Using the artificial training datasets, the DNN was then trained to establish a relation between the depositional characteristics and the

and the flow conditions. The post-trained DNN model was ready to predict flow conditions from the tsunami deposits after the performance of the trained DNN was verified using test data sets. The 1-D cubic interpolation was applied to the field data sets of Phra Thong island to fit the dataset to model grids. Finally, this DNN inverse model was applied to the field data sets from the Phra Thong island, Thailand to reconstruct the flow conditions of 2004 Indian Ocean tsunami. We also used the reconstructed flow conditions to estimate the spatial distribution of the volume per unit area and grain-size composition from Phra Thong island and compare the distribution with the measured data. Our inverse model was already validated to be effective for 2011 Tohoku-oki tsunami deposits distributed in Sendai Plain. In case of Phra Thong island, we validated the results by the field measurements of the tsunami flow depth. Also, the estimated thickness and grain size distribution of tsunami deposits were compared with the actual measurements. Our inverse analysis results could be used for designing future tsunami hazard assessments and disaster mitigation strategies in Thailand.

10. P-3, Line 83, If you mention this, then one would expect to see grain-size variations and deposit thickness data.

    RE: Thank you for the suggestion. We will add a diagram which represents grain-size distribution and thickness data.

11. Line 105, This is a big simplification...

    RE: Thank you for the comment. This simplification was done by Naruse and Abe, 2017 and this was further used by Mitra et al., 2020. For the details of the simplifications and step by step procedures, please refer to Naruse and Abe, 2017.

12. P-5, Line 115, What one expects from an inverse model (like you are suggesting in the title) is that you depart from field data and use it to reconstruct the physical parameters of the tsunami. This is not what you are explaining here. You use the forward model to produce tsunami hydraulic features and then based on the neural network you check where the data produced is nore robust and assume the values produced. I was expecting a regression from deposits to flow characteristics. It does not seem to be the case here.

    RE: Thank you for the comment. Yes, we are using the field data to reconstruct the physical parameters using the DNN inverse model. The forward model is used to generate artificial spatial grain size distribution and volume per unit area data of tsunami deposits which was used to train inverse model. We checked the robustness and precision of the inverse model using the artificial test dataset, and then used the produced inverse model to reconstruct the flow conditions of the tsunami from the actual deposit.

13. P-5, Line 119, I understand it but you need to clearly show where is the data coming from. can you please add a map and grain-size data.

RE: We assume that this comment is related to comment number 10. Please refer to the reply in comment number 10.

14. P-5, Line 134, Why 0 to 2% concentration? I do not understand where this value comes from? Is it random? From observations? Where? All the data I have access to suggest (much of it publiched by several authors) much higher sediment concentrations.

RE: Thank you for the comment. This range was considered from Goto et al., (2018) sediment concentration analysis which indicates that the total sediment concentration for tsunami deposits is usually around 2%, and therefore the concentration for each grain size class seems unlikely to exceed 2%. In response to your comment we will modify our text by adding this reference. Please refer to the reply of comment 2 where we have written the modified text.

15. P-5, Line 145, This needs to be explained earlier and more clearly.

RE: Thank you for the suggestion. We will revise the text in the revised manuscript.

16. P-6, Line 157, I understand the seminal character of the work by Mitra et al (2020) and the need to many citations along the manuscript but is some cases I believe this could be avoided.

RE: Thank you for the suggestion. We will remove the sentence from the revised manuscript.

17. P-7, Line 161, What were their charateristics?

RE: The characteristics of artificial data sets were depositional characteristics such as volume per unit area and grain-size distribution.

In response to your comment we will revise the text in the revised manuscript.

18. P-7, Line 165, And between 2000 and 3000 what is the gain?

RE: The loss function was optimized already to its lowest value, the calculations continue up to 3000 epochs, the progress will remain the same. There will be no change as it already converged before 2000 epochs.

19. P-7, Line 169, this is obvious. No need to add it.

RE: Thank you for the suggestion. We will remove it from the revised manuscript.

20. P-8, Line 188, Please add a aerial image with the sampling point clearly marked.

RE: We agree with the reviewer. We will add a google earth image with the marked sampling points.

21. P-8, Line 191, add reference. Heights were reported or modeled?

RE: Thank you for the suggestion. We will add the reference Fujino et al., (2010) as reference. Heights were reported from the water mark at the study area. We will modify the text by mentioning this in our revised manuscript.

22. P-8, Line 192, the nearshore area?

RE: Yes. The area was nearshore.

23. P-8, Line 193, thin and finer?

RE: The deposits became thinner and finer in the landward direction. We will revise our text by adding both phenomena.

24. P-8. Line 196, along the analyzed profile, right? Otherwise it contradicts what is stated above regarding maximum thickness of the deposit.

RE: Yes, all the thickness mentioned here is along the analyzed profile.

25. P-8, P 198, This clearly needs to be provided here as well. This is pivotal for the validation and must be displayed in this manuscript as well.

RE: We assume that this comment is related to comment number 10. Please refer to the reply in comment number 10.

26. P-10, Line 200, Not sure if I understood this last sentence correctly. Please provide fruther information. So, you stopped using field data?

RE: Thank you for the comment. We have done subsampling test to check the effect of irregularly spaced field datasets on the accuracy of the inversion. The details on the subsampling procedure is given in Mitra et al. (2020).
We agree that we should add some basic information or the add the reference in the text. We will modify the text in our revised manuscript.

27. P-14, Line 222, Please clarify how you obtained these values.

RE: Thank you for the comment. These values are obtained from table 1. We will add clarifications on how we obtained this data in the revised manuscript.

28. P-14, Line 227, Discussion needs to be rewritten.

RE: Thank you for the comment. We hope to modify our discussion accordingly.

29. P-14P, Line 232, So model was not validated against field data?

RE: The model reconstructed the tsunami flow characteristics from the actual tsunami deposit, and the predicted results were validated with the field measurements of flow depths. This model was trained with artificial datasets of tsunami deposits and validated with the field data. The present model is estimating fair results using inexpensive artificial data for training of the neural network and avoiding the difficulties to gather large amounts of datasets of tsunami deposits with in-situ measurements of flow velocity and depth. For tsunami deposit dataset in terms of grain size distribution perpendicular to coastline is not easily obtainable and sometimes good datasets are difficult to obtain due to obstructions in field areas, and measurements of flow hydraulic parameters such as velocity were quite rare. Therefore, training the model with the real measurements may not be most viable option in terms of expense and performance. Indeed, all of previous studies on inverse analysis were not developed from the relationship between measurements of deposits and flow parameters because it is practically impossible. Instead, their models were depended on the simplified hydraulic modeling of tsunamis. Even if we can train model with smaller number of training data obtained by measurements in the field, the model tends to overfit which results into poor performance of the model on the observed value.
To this end, we employed very different approach, which uses the calculation of results in the forward model as the training datasets.

30. P-14, Line 239, when you mention true values are you referring to field data?

RE: We agree that this true value might create confusion. We referred the true values as the test dataset from artificial data that determines the performance of the model. We will clarify this terminology in the revised manuscript.

31. P-16, Figure 9, For me this is the key findings of the manuscript and are somewhat lost in the structure. This is a key figure that should be move further up. It will also important to acess raw/original grain-size curves retrieved from the field by Fujino et al. (2010)

RE: Thank you for the comment. We considered shifting this diagram in the upper part of the section.

32. P-17, Line 257, This sub-section is crucial and is poorly explained above and consequently also here. The authors should make an effort to clarify the field validation. I read the manuscript 3 times and struggle to fully understand what field data you used. And if it is an regressive exercise why you do not use thickness and grain-size curve or D50. This needs to be clear for the reader from the start of the manuscript.

RE: Thank you for the comment. We understand that there must be lack of clarification in the description of validation regarding the application of the field data. We hope to explain in more details in the revised manuscript.

33. P-17, Line 266, measured or modeled?

RE: The flow heights were measured using water mark in the field. We will add this detail in the revised manuscript.

34. P-17, Line 269, observed or modeled? Did someone measure flow height at these specific points or these values are the result of forward modeling exercises? I suspect it is the latter. So change the text, please to be accurate.

RE: Yes, the mentioned groups measured the flow heights at specific locations. The details are available in the link, http://www.nda.ac.jp/ fujima/TMD/fujicom.html.
In response to your comment we will unify the terminologies of measured flow heights.

35. P-17, Line 274, But you used Fujino et al., 2010, right?

RE: We applied the dataset of volume per unit area and grain size distribution in our model and the model results were validated with several reported or measured values by different researchers. Here, the distance 400 m indicates the locations of the measured value by Jankaew et al. (2011, 2008) from our study area.

36. P-18, Line 296, But there is a wide range of values presented even in this manuscript introduction and study area?!?!

RE: The DNN inverse model estimates only a single value. We have added the uncertainty or error estimations. We used a range of artificial values to generate artificial dataset of depositional characteristics in the forward model. Using that artificial data sets the inverse model was trained and the final predicted result was a single value for one sampling window size. For example, we used 1700 sampling window size to reconstruct the single values for each parameter of flow conditions.

37. P-18, Line, 307, Please explain this idea better and in greater detail.

RE: Thank you for your comment. We will modify the text and some further details in the last paragraph.

38. P-18, Line 310, Where have you provided this information? Where is the detailed variation of tsunami deposit thickness variation in the field and its comparison with the model results?

RE: Figure 9 shows the tsunami deposit volume per unit area and its comparison with the model results.

39. P-18, Line 315, Please rewrite the Conclusions.

RE: Thank you for the suggestion. We agree that we should rewrite the conclusion in more details. We will do so in our revised manuscript.