

# Response Letter 1

Manuscript: "Identifying the drivers of private flood precautionary measures in Ho Chi Minh City, Vietnam" by Thulasi Vishwanath Harish, Nivedita Sairam, Liang Emlyn Yang, Matthias Garschagen, and Heidi Kreibich

**Our Response R: *In Italics***

*We would like to reviewer for taking the time and effort to provide comprehensive feedback on our manuscript.*

General Comments:

**Incorporation of flood risk mitigation measures taken by the residents is an important part of flood risk analysis/management. This study aims to identify the drivers of the private sector to implement these measures in Ho Chi Minh City, Vietnam by analysing the survey data collected from 1000 at-risk households. For this purpose, the researchers used 'Protection Motivation Theory' in combination with the 'Transtheoretical Model' by accounting for both non-structural and structural measures based on both 'Proactive' and 'Reactive' behaviours. The article includes some interesting contents that are substantially practical in the real-world flood risk analysis. However, there lies a number of fundamental concerns, mostly related to the organisation and the flow of the information on the methodology and application.**

*R: We thank the reviewer for taking their time to provide a comprehensive review of our manuscript. We have restructured the manuscript in accordance with the reviewer's comments to improve our manuscript. In addition, we have provided a detailed explanation to our approach in the response letter and the manuscript accordingly. The lines cited from the manuscript are provided in quotes.*

Major Comments:

- **It would appear that the majority of the contents are parts of a larger document that are put together without a decent amount of cohesion and linearity. This significantly disturbs the reader to follow the flow of the article and understand the novelty of the proposed method as well as the usefulness of its output in the context of flood risk management. More specifically:**

1. A bulk body of information in the ‘Introduction’ section constitutes the historical data on the flood events (lines 21-37), introduction to Flood Risk Management (lines 39-48) and the private precautionary measures (lines 49-65). Though important elements (and probably important for a thesis or dissertation), they are not specific to the research novelty and the method that the authors employed to “Identify the drivers of the private flood precautionary measures in Ho Chi Minh City, Vietnam”. However, Section 2, ‘Study area – Ho Chi Minh City’, appears to include the introductory information that the reader should obtain from an introduction section which is more focused on research itself. Also, the introductory information in Section 3 could aim the authors to rewrite the introduction section.

*R: Thank you for the comment. Yes, the study was performed as a part of a Master’s thesis. We have now improved the structure and coherence of the manuscript by removing some parts from the introduction that are not relevant to the study. The study area is now included in the introduction.*

2. Section 3, which is expected to convey the information on the ‘material and method’ in the research is not self-explanatory of what is specific to the research considering the aim and objectives. The authors could use the framework introductory information, explained in Section 2 (lines 99-124), in the rewritten ‘Introduction’ section and focus more on the materials and methods in the new section (such as lines 127-204).

*R: We have incorporated the suggested changes. The introduction to framework is now included in the introduction. Materials and Methods now includes a sub-section on the questionnaire survey, application of the PMT-TTM framework relevant to the study and the description of data-driven ML models.*

3. Section 4, Results and discussion, does not respect the standard of a decent academic article and therefore should be rewritten to comply those requirements (authors could use other examples published in NHESS). More specifically, the section and the subsections lack introductory sentences to acquaint the readers with appropriate preparatory information for what sort of outputs will be discussed with respect to a specific purpose in the coming subsections. Take the instance of line 206, which jumps directly to the specific outputs without any preparation to answer: ‘what’, ‘why’ and ‘how’ with respect to what was said before in the ‘materials and method’ section. Also, this section includes a large amount of discussions on the previous research other than specifying the key findings of the present research.

*R: We have improved the ‘Results and discussion’ section by including appropriate preparatory information, as suggested. Please see answers to specific comments where this is explained in detail.*

4. **Section 5, conclusion, should include the key findings of the research alongside the brief summary of the research. It should also include specific limitations in more details and suggestions for future research. It would also help the quality of the conclusion to provide information on the reproducibility of such outputs for other place/countries.**

*R: The conclusion section has been revised to include the key findings and the summary of the research along with the scope for further improvements.*

- **The article does not weigh the privilege of the proposed research over other existing methods. For example, questionnaire surveys. In simple words, the article does not highlight the importance of using ‘Protection Motivation Theory’ and ‘Transtheoretical Model’ for conducting such research.**

*R: The study uses a combined protection motivation theory - trans-theoretical model as the analysis framework and uses data from questionnaire surveys to quantitatively derive the drivers of private precautionary measures. The advantage of this methodology is now explained in the introduction “The privilege of the research is that the framework provides an understanding of the cognitive processes - threat and coping appraisal and the decision stages involved in implementing private precaution. The questionnaire data provides case study relevant inferences which has the potential to be used in risk communication and management”.*

Technical/Minor Comments:

- **The ‘Abstract’ should be revised so as to clearly elaborate on the method, what has been specifically analysed from the survey datasets, and what are the key findings of the research and what do they imply/show. Especially line 13 onwards, the flow of information does not seem to be correct; therefore, makes it difficult to understand. Also, avoid using long sentences as short ones would help the readability of the abstract and all your work.**

*R: Thank you for the comment. Abstract is revised to include details on the methods, key findings and implications.*

*“Private flood precautionary measures have proven to reduce flood damage effectively. Integration of these measures into flood response systems can improve flood risk management in highly vulnerable areas such as Ho Chi Minh City (HCMC). Since uptake of such measures is voluntary, it is important to know what drives householders to implement precautionary measures. In this study, we developed a framework representing the uptake of private precautionary measures based on Protection Motivation Theory and Transtheoretical Model. Using empirical survey data collected from 1000 flood prone households in HCMC, we*

*implemented ridge and elastic net regression to identify the drivers of private precaution. The measures were classified into structural measures and non-structural measures based on whether structural changes to the building were required. The household profiles were classified into proactive and reactive households based on whether their decision to reduce risk (i.e. uptake precautionary measures) depended on their past flood experiences. The data-driven model revealed that the level of education, the degree of belief in the government to implement regional flood protection measures and the degree of belief that in case of flooding, one has to deal with the consequences of flooding by themselves positively influence the proactive uptake of non-structural measures. Among the households that experienced flooding in the past, the uptake of measures was found to be driven by the severity of the experienced damage. On the other hand, perceiving high severity of future flood impacts was found to negatively influence the implementation of structural flood precautionary measures. These results recommend that efforts to communicate flood risk awareness along with the knowledge of private precaution in HCMC should consider the socio-economic characteristics of the household, their past flood experience and perception of future flood risk.”*

- The ‘Introduction’ section has many repetitions, which could be made more concise with respect to the aim and objectives of the research. For example, doesn’t line 24-26 convey similar meanings to the previous two lines?

*R: Thanks for the comment. These lines were extracted while revising ‘Introduction’ and other repetitions are now removed.*

- In-text citations does not seem to follow a uniform template. In some instances, a comma is used after et al. while in the others there is none. Compare for example, line 31, (Nguyen, et al. 2021(b)) and line 36, (Cao et al., 2021).

*R: In-text citations are rechecked for uniformity in the revised manuscript.*

- Line 31-33, ‘Developing countries … ‘could be more specific. For example, by specifying ‘What limited capacity?’

*R: The suggestion has been inserted in the revised manuscript.*

*“Developing countries are more severely impacted by natural disasters such as flood events due to their limited economic and infrastructure capacity to lessen the effects of disaster and this hinders their economic growth (Botzen, et al., 2019(a); Parker, 2006).”*

- Line 36. The authors have not reviewed any research yet; therefore, it would be better to provide some more explanatory information on the physical and environmental drivers of flood risk before providing such conclusion.

*R: Thank you for the comment. We referred to the literature review by Nguyen et al., 2021(a) (<https://doi.org/10.1111/jfr3.12689>) concludes that most studies tend to prioritize physical and*

*environmental drivers. However, Line 36 is removed to maintain coherent flow in the modified introduction of the revised manuscript.*

- **Line 37. “To counteract the trend of increasing flood risk due to global change, improved flood risk management is necessary.”** is a trivial piece of information and sounds unnecessary as it has neither related to what has been said before nor has been specifically in line with the flow and aim and objective of your research.

*R: We agree. The sentence has been extracted.*

- **Line 42 is a good place to explain about proactive measures in the context of the implementation of flood risk management strategies.**

*R: Thanks for the comment. The word ‘proactive’ has a specific definition in the context of this research – i.e. we refer to households’ uptake of private precaution without flood experience as a driver as proactive decision-making. Therefore, we would like to explain this term in ‘Data and Methods’ section where we explain the decision stages.*

- **Lines 42-45: the authors could elaborate more. Do you mean: because the flood hazard changes rapidly in urban areas amongst the household units, implementation of the conventional large-scale flood protection measures, such as dikes and retention basins, is challenging?**

*R: Yes, the explanation is correct. We have provided more explanation in the revised manuscript – “Large-scale structures are designed to provide protection against flood events below a specific return period but increasing intensity and frequency of flood events is exposing growing urban areas to flooding making these structures less reliable”*

- **Lines 45-48: is not comprehensible as the previous paragraph is not structured based on the research scope.**

*R: After providing explanation to the previous comment, we believe this sentence is now comprehensible - “Since large-scale protection structures are designed to a specific return period, we need to complement these structures with additional measures such as land-use planning, private precautionary measures and insurance to cope with residual flood damage.”*

- **Line 49: use 'Private precautionary measures have demonstrated to be effective in reducing flood damage.' instead.**

*R: We have changed the sentence accordingly.*

- **Line 53: use “There is a knowledge gap in ...” instead.**

- R: We have changed the sentence accordingly.
- Line 60: “Experiencing repeated flooding can change this attitude (Bubeck et al., 2018; Chinh et al., 2016).” Using such sentences from other sources requires mentioning further backup from other research. For instance, what did they specifically conclude in their research? Using such assertion for your purpose in the introduction does not provide the reader with the required clarity.

R: Thanks for the comment. The statement has been extracted from the introduction in the revised manuscript.

- Line 63: “These insights can guide the design of targeted risk communication campaigns and incentives to improve flood preparedness” The authors should elaborate more on such sentences.

R: Thank you! In the revised ‘Introduction’, this is improves - “The privilege of the research is that the framework provides an understanding of the cognitive processes - threat and coping appraisal and the decision stages involved in implementing private precaution. The questionnaire data provides case study relevant inferences on the propensity of households to uptake private precautionary measures. This has the potential to develop household profiles that guide the design of targeted risk communication campaigns and incentives to improve flood preparedness”

- Lines 49-71: This paragraph should be more specific and concise to fit the purpose of this study. The authors should provide a summary of what has been done in the present research as well as its basis and novelty.

R: Thank you. The revised introduction is restructured and improved based on the given suggestions.

- Line 76: add reference for “The city's population is expected to grow even faster in the coming years.” Also, it is good to mention the population growth rate.

R: The reference has been added in the revised manuscript - (Asian Development Bank, 2010).

- Line 82: The flood risk is exacerbated by climate change, ongoing urbanization, increasing population and infrastructure density leading to a higher proportion of sealed surfaces.” Is a repetition. The authors should consider removing the repetitive sentences in the revised/rewritten manuscripts, especially in the ‘Introduction’ and ‘Materials and Method’ sections.

R: We have removed the redundant sentences.

- Line 94-96: as last lines of your subsection should conclude your discussion on the issue. Please consider to follow a more linear approach in providing the reader with the required information before they reach the next section of your manuscript.

R: *This section has been moved to introduction (along with the rest of the case study explanation) and rewritten accordingly.*

- As mentioned above, the majority of Section 3 could be used in the revised ‘Introduction’ sentence and this section should be more concise and specific to the researcher’s own work and method.

R: *We agree and have implemented the changes.*

- Line 119: use “The present approach” or “The proposed framework” instead of “This framework”.

R: *Thanks! It is altered in the revised manuscript.*

- Line 129: “The survey collected 1000 valid responses from local households who suffered from floods in the recent 10 years.“ here the authors are expected to mention the representation percentage of the selected number of household with respect to the total population (9 million + 2 million?). And also, how this rate would influence the validity of the research findings?

R: *In this study, we invested a large amount of effort in the selection of survey areas (see figure 1 in the revised manuscript) than specific households. The survey areas were identified together with local stakeholders and research partners in HCMC. The chosen areas were diverse and represent various heterogenous urban characteristics. They were all prone to frequent floods. The sample is representative of the diverse characteristics of households in the growing urban regions of HCMC that are prone to frequent flooding. Owing to the diversity across the surveyed areas, changes of the specific number of surveyed households, e.g. from 1000 to 1500 or 800 is supposed to influence the research findings at a very limited level. We have included this point into the updated manuscript.*

*This is now included in the revised manuscript –*

*“The empirical data used in the study was obtained from a structured household survey in selected districts of HCMC during September - October 2020. A total of 8 wards in 4 districts were surveyed which includes Binh Thanh, District 8, Binh Tan, and Nha Be. The survey collected 1000 valid responses from local households which suffered from floods in the last 10 years. The questions were drafted based on expert knowledge from flood risk researchers, social scientists and local stakeholders in HCMC. The survey locations were established in*

*order to cover a broad range of socio-economic profiles and flood types such as tidal, fluvial, pluvial and compound flooding in the city. A survey pre-test involving 60 households from three districts (Binh Tan, District 7 and District 2) was run in December 2019 in order to test the validity of the questionnaire. The questionnaire was revised, based on the responses from the pre-test. The questionnaire covered aspects concerning two past flood events experienced by the households - the most recent and the most serious event in the last 10 years. The questions pertained to the hazard and damages suffered by the households, implementation of precautionary measures, early warning quality and lead time, household risk perception and household's socio-economic profile."*

- **Lines 131-132: Could be more specific. For example, what class of socio-economic profiles and what types of flood types have been investigated?**

*R: We didn't intend to select special household classes of socio-economic profiles. In contrast, based on expert advice from stakeholders and researchers in Vietnam, we invested on covering diverse urban neighbourhoods in HCMC. At the household level, the only criteria in the survey was that "the household must have experienced floods in the last 10 years". The flood types included pluvial, fluvial, tidal and compound events.*

- **Line 152: use "Each precautionary measure is categorised into ...".**

*R: Changed accordingly*

- **Figure 2. Is this figure necessary? The authors could use a simple table instead.**

*R: Figure replaced by 2X2 matrix*

- **Line 164: the authors should explain the reason that the lasso and elastic net regression models lead to identification of drivers. And in-detail explanation is required here.**

*R: In detail explanation is provided as suggested in the revised manuscript.*

*"To identify the important explanatory variables that influence the decision of a particular household group to uptake private precaution, lasso and elastic net regression models are applied. These models determine the relationship between the explanatory variables representing one of the aspects in the PMT-TTM framework and response variables indicating if a precautionary measure was implemented or not, for each household group. The response variables follow binomial distribution.*

*Lasso regression determines the extent of influence by an explanatory variable on the response variable by imposing  $\lambda$  times L1 penalty on the residual sum of squares to compute the lasso estimate as defined by Eq. (1) (Hastie et al., 2008):*

$$\hat{\beta}^{lasso} = \operatorname{argmin} \beta \left\{ \frac{1}{2} \sum_{i=1}^N (y_i - \beta_0 - \sum_{j=1}^p x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^p |\beta_j| \right\} \quad (1)$$

Here,  $x$  represents explanatory variables,  $y$  is the response variable,  $\beta_0$  is the intercept,  $\beta$  represents regression coefficients of explanatory variables,  $p$  is the number of input explanatory variables, and  $N$  is the number of observations or households interviewed ( $N=1000$ ). L1 penalty is  $\sum_{j=1}^p |\beta_j|$  while  $\lambda \geq 0$  is a complexity parameter that controls the amount of regression coefficient shrinkage and is determined by cross validation. Larger the value of  $\lambda$ , greater is the shrinkage (Hastie et al., 2008). Lasso regression performs variable selection while maintaining the stability by imposing a penalty on the size of regression coefficients (Tibshirani, 1996) and shrinking it towards zero when there is low correlation between explanatory variable and response variable. The nature of this constraint tends to produce some coefficients that are exactly equal to zero and eliminates the explanatory variables corresponding to these coefficients (Tibshirani, 1996). However, when a number of explanatory variables ( $p$ ) is greater than the number of observations ( $N$ ), only  $N$  variables are selected before lasso saturates and when a group of variables have high pairwise correlation, then lasso randomly selects one variable from the group. Therefore,

A naive elastic net regression is introduced as illustrated in Eq. (2) (Zou and Hastie, 2005).

$$\hat{\beta}^{naive\ elastic\ net} = \operatorname{argmin} \beta \left\{ \sum_{i=1}^N (y_i - \beta_0 - \sum_{j=1}^p x_{ij}\beta_j)^2 + \lambda_2 \sum_{j=1}^p \beta_j^2 + \lambda_1 \sum_{j=1}^p |\beta_j| \right\} \quad (2)$$

$$\text{Were, } \alpha = \frac{\lambda_2}{\lambda_1 + \lambda_2}$$

It possesses the characteristic of lasso regression which uses the L1 penalty,  $\sum_{j=1}^p |\beta_j|$ , and ridge regression which uses the L2 penalty,  $\sum_{j=1}^p \beta_j^2$ . It combines the advantages of L1 Lasso penalty which performs automatic variable selection and L2 ridge penalty which encourages grouped selection by shrinking together the coefficients of correlated explanatory variables (Hastie et al., 2008). Hyperparameter  $\alpha$  estimates the contribution of L1 and L2 penalty by assigning a value between 0 and 1. However, Eq. (2) was unable to perform satisfactorily as its solution path incurred double-shrinkage and did not produce optimal variance-bias trade-off. Rescaling naive elastic net equation by  $(1+\lambda_2)$  as shown in Eq. (3) automatically achieved optimality and is known as elastic net regression (Zou and Hastie, 2005).

$$\beta^{elastic\ net} = (1+\lambda_2) \hat{\beta}^{naive\ elastic\ net} \quad (3)"$$

- The authors should explain the notations in all the equations and avoid explaining the repeating ones. This can be done by providing a couple of lines below each equation/formula.

R: Notations are explained in the revised manuscript. For example, see the response for the previous comment

- Date shown in Figure 4: where did the authors derive/obtain the ‘Implementation cost of the private precautionary measures’? If it is a part of the present research, there has to be some explanation. If not, the authors should provide information on how they obtain them. Also consider it in the further discussion.

*R: The costs are part of the present research. They were a part of the survey questionnaire. Explanation to Figure 4 is included for better understanding – “Figure 4 exhibits the distribution of cost incurred by a household to implement a precautionary measure. The costs were obtained during the survey from the households that implemented the measures.”*

- Lines 240-245: Study worth analysing the difference in the socio-economical drivers between the countries that influence households to take flood mitigation measures at individual levels. What are the differences and how they might change according to each country socio-economical driver?

*R: Thank you for the comment. We attribute the differences in choice of precautionary measures to structural differences in houses – frequent renovations leading to elevation being the most prevalent measure and flood situation – frequent flooding (nuisance flooding) driving the implementation of permanent structural measures. We removed the reference to socio-economic drivers – “Majority of the respondents in this study have exclusively only elevated their house which is often built elevated or are elevated during renovations. The process to elevate can be done to the entire building or only a new elevated ground floor can be constructed within the building (FEMA, 2007). Households in HCMC often make structural renovations resulting in ‘elevation’ being more prevalent. Other measures such as installation of flood protection systems and usage of water-resistant materials, though found prevalent in other regions, are not common in HCMC. This might be due to lack of knowledge or lack of support to increase responsibility among households to implement other private measures (Bubeck et al., 2012; Chinh et al., 2016). Water barriers are convenient to implement without the need to make permanent structural changes to the building and they prevent flood water from entering the house. Wet proofing also prevents the contact of valuable items with flood water. However, these measures can be effectively implemented only when flood warning is given in advance. Highest number of respondents have structurally elevated their houses only after experiencing serious and recent flood events (Figure 3) despite the high cost of implementation because of its effectiveness in preventing the floodwater from reaching the living area. However, a study by Koerth et al. (2013) states structural measures as the least preferred flood precautionary measure by the households in Denmark and Germany due to their high costs. Similar study conducted by Bubeck et al. (2018) in Germany and France, reports the requirement of policies to encourage structural flood-proofing. Thus, the choice of private precautionary measures by households in HCMC starkly differs from the prevalent measures in western Europe. HCMC’s mitigation choices are driven by frequent floods, inundations occur almost during every rainy season, so that people learn to live with floods.”*

- **Section 4.2 (starting from line 245):** The variables discussed here are not mentioned/explained previously in the article, therefor the reader is not familiar with these terms. The authors can make a table and explain each before the reader reaches to Section 4.2.

*R: All the variables are described in Table 3 and Appendix A. The reader is also directed towards these tables. However, we understand the lack of coherence and have restructured the section in order to get acquainted with the variable description in the Table before reading them in the text.*

- **Lines 251-253:** Good to know what Table 3 has. It should be explained before discussion on the results. Please do not mix the discussion of Table 3 with Figure 5. The authors should first explain Table 3, then discuss figure 5. The authors should ensure that they explain about the variables before jumping into the discussions. It can be in a few lines in the introductory paragraph of this section explaining what has been studied with what aim and how.

*R: As previously mentioned, this section has been restructured as suggested and Table 3 is introduced well in advance.*

- **Lines 253-258:** the meaning of these sentences are not clear. For example, what is the difference between 'house damage' and 'house impact'? The authors should explain 'house damage' and 'house impact' before reaching here to give the readers an idea of what they mean.

*R: The sentence is more comprehensible after reading the variable description column in Table 3 which has now been moved to the starting of the section after restructuring. 'House Damage' variable represents the damage already experienced by a household from past flood experience. 'House Impact' variables depicts the level of damage to the house anticipated by the household due to future flood events.*

- **Line 294: use comma after 'Next'.**

*R: comma added to the text.*

- **Lines 297-299:** There is no need for mentioning the findings of the previous research in your discussion if they are not related to the results' discussion.

*R: We would like to retain the examples from previous research. It is interesting to compare our findings with previous research and provide a perspective on case-study relevant differences in drivers of private precaution.*

- **Lines 301-305:** require further elaboration.

*R: Thanks for the comment. We are unfortunately not able to infer more on the drivers for structural proactive measures. This is also attributed to data limitation. We have included this in the manuscript - "Methodologically, there is a lack of variability among the structural proactive group of households, only 264 out of 1000 households have proactively implemented a structural measure while there are 30 explanatory variables. Therefore, each explanatory variable is supported by less than 10 households with lower frequency outcome. Therefore, no important variables were identified."*

- Lines 306-319: This is not a concluding paragraph for the discussion section. The authors can clearly explain what their discussion suggests with respect to the aim and objective of the research.**

*R: Concluding paragraph of the section has been revised - "the long-term improvement in precautions taken by commune authorities and households appears as inseparable components of the integrated flood risk mitigation strategy (Nguyen, et al., 2021(b)). Findings in this study are especially relevant in many fast-developing Asian cities due to their similar features regarding high population density, strong social network and spontaneous individual coping measures. At household level, certain indicators including education, income and place attachment can be empirically valid for taking active actions (Ji, et al., 2021). Both wet-proofing and dry-proofing measures are widely adopted as the most common precautionary measures (Lasage, et al., 2014), along with ground elevation, foundation strengthening, using reinforced materials, precautionary savings, in various countries, covering coastal cities (Du, et al., 2020) as well as inland regions (Shah, et al., 2017). However, findings confirmed that the high level of uptake of precautionary measures was dependent on distance, household composition, income, occupation of the household and social network type (Okayo, et al., 2015). In addition, households may be better aware of the limitation of public flood protection through their precautionary behavior, and thus found which measures can reduce most of the flood damages (Kuo, 2016). Therefore, it would be interesting to do a comprehensive analysis on the motivating factors that can be used to move reactive households to the proactive group. To further explore this specific theme, empirical data on the cost and benefits of flood precautionary measures would be paramount."*

- In 'Competing interests', Is this necessary to mention that one of the authors is an executive editor at Natural Hazards and Earth System Sciences (NHESS) journal? Please do check it with the editor. Also, make sure that the Appendix is located in a correct place.**

*R: We have removed the statement mentioning one of our authors is an executive director at NHESS under 'Competing interests' section. According to the Manuscript composition ([NHESS - Submission \(natural-hazards-and-earth-system-sciences.net\)](http://natural-hazards-and-earth-system-sciences.net)) Appendix is rightly placed after conclusion.*

- **Titles of the sections and subsections should be more informative elaborating on their contents by also preserving the linearity in the revised manuscript.**

*R: Revised as suggested*

- **Also consider using more informative caption for the figures and the table 1.**

*R: Revised as suggested.*