Response to Peer Review Comments

Referee #1:

Major criticisms

With respect to the two existing models, neither the applicability nor transferability of the model are analysed. To this aim, I expected authors to analyse whether the hazard and vulnerability features considered in the development of the models fit with Australian conditions; whether vulnerability parameters of the models are significant or not in the study area, etc.
The authors simply evaluate the accuracy of damage estimates (supplied by the two models) by a comparison with empirical data.

- We appreciate the comment. Based on the reviewer suggestion, the objective of study will be revised. Also, a new dataset has been provided by the reconstruction authority of Queensland. Therefore, the newly derived model will be calibrated with the historic data collected from Bundaberg region. Afterwards, performance of all applied models, for estimation of structural damage in the Australian study area, will be checked by the dataset collected from Maranoa region.

As regards the new model, I understand that it was calibrated and validated with the same dataset which of course implies good performances. The comparison of the performance of the three models by using a different set of damage data, would be more reasonable. In general, the description of the new model requires some clarifications (see below)

- We agree with the reviewer that the previous comparison could not guarantee performance and validation of the newly derived function. In this way, as explained above, two different datasets will be used for the calibration and validation of the new model.

Also the performance evaluation (section 6) suffers of several problems. See below.

- It has been discussed below.

The present organisation of contents creates confusion in the reader. Sometimes linked concepts are discussed in distinct sections; some concepts are instead not in the right section.
I would suggest to change the objective of the paper that should be "presenting a new model for Australia and comparing it with existing methodologies (one local and one from US)".

- We are grateful for your suggestion. The objective of paper will be revised based on your suggestion.

To this aim I would suggest to change the organisation of contents in this way:
- remove section 3 and replace it with a section describing in detail the new model. The description of the GA method and USACE method can be directly included in section on results.
- Section 5 should focus on the derivation of the model, including the statistical analysis.
- Section 6 should focus only on models comparison.

- We appreciate your suggestion. We will revise the order of content based on your suggestions.

Specific (minor and major) comments

Section 1

Pg 2 line 24 "in recent decades, the probability of flood and the values of exposed properties have increased exponentially" → This is not totally correct and is not supported by the articles cited. An increase in flood damages has been observed (as suggested by the authors quoted). However, a debate is ongoing about whether such a trend is due to an increase in hazard frequency and intensity or to a better damage recording. I agree that exposed properties increased significantly.

- It will be revised as: "In recent decades, the flood risk due to climate change and the growth in value and vulnerability of exposed properties has been increasing exponentially (Elmer et al., 2012)."

Pg 3 line 1 "*damage assessment in terms of mitigating the probability of expected losses*" → Why "in terms of*"? Do authors mean "damage assessment as the base for/ in order to mitigate the probability of expected losses"?*

- Corrected as: "in order to"

Pg 3 line 27 "On the other hand, flood actions on buildings could be related to a variety of hydraulic factors such as lateral pressure, velocity, duration, debris, erosion, and the chemical effects of water. But most of the models for direct impact estimation consider only the depth of water as the main characteristic of flood". →The concept is not "concluded". Which is the consequence of this? I guess model uncertainty. If this is the case, the role of vulnerability factors should be discussed as well, because they can be even more influencing.

- We appreciate this comment. It will be revised as: "Flood actions could be related to a variety of hydraulic factors such as lateral pressure, velocity, duration, debris, erosion, and the chemical effects of water. However, water depth is identified as the most dominate influencing factor on residential buildings flood damage (Merz et al., 2010; Thieken et al., 2005; Cammerer et al., 2013)"

Pg 4 line 14 "Although the simplicity of stage damage functions is the main reason for their common usage, neglecting some influencing aspects due to a lack of real damage data will raise the level of uncertainty". → this is actually the conclusion of the above concept. But it appears some lines below. Suggestion: the role of influencing factors (or explicative variable I would say) on model uncertainty is not in the scope of the paper. There are not analyses/evaluations on the role of considered and neglected factors on model accuracy. In order to avoid confusion in the reader, I would delete this discussion from the paper or, at least, I would specify that the point is not handled in the following.

- It will be revised as: "The simplicity of stage damage functions is the main reason for their common usage. However, using them in different geographical regions without evaluating their performance would raise the level of uncertainty in flood risk assessment (Cammerer et al., 2013)."

Section 2

Pg 4 line 19 *"In addition to the Rapid Appraisal Method (RAM), which is an averaging methodology for damage estimation, there are a lot of depth–damage curves for flood loss assessment in*

Australia. The RAM is a simplified method for flood damage estimation in the absence of data required for using depth–damage curves. This method considers mean unit values of damage for all buildings in the inundated area. Although RAM is useful for early assessment of the magnitude of damage, the results are considerably inaccurate (Barton et al., 2003)". → *I do not understand why the RAM method is explained as it is not subsequently used. Ok in the introduction, not in the background section.*

- We explained the RAM methodology in the background, because we wanted to discuss the state-of-the-art nature and knowledge gaps of Australian methods that motivate us to propose a new method. However, as you have mentioned correctly, this short explanation will be replaced to the introduction part.

Pg. 5 line 2. "*Gathering data from one actual flood event and using it as a guide for future events in a new area of study, or even in the area of origin, requires a complicated process of extrapolation (Gissing and Blong, 2004; Smith, 1994)" .* → quite general, please explain

- We will add an extra explanation as: "in other words, extrapolation of empirical damage curves to different regions is difficult due to differences in the level of precaution and differences in building characteristics (Barton et al., 2003)."

Pg. 5 line 6 *"Valuation surveys refer to the value and elevation of all assembly items and contents that are located above the basement"* → *Not clear, please specify more*

- We will add an extra explanation as: "this means that by using valuation surveys, an average distribution of building fabric in the height of the structures would be extracted."

Pg. 5 line 27 *"ANUFLOOD"* → Not all readers know such a model. Please introduce it or, at least, supply references.

- We appreciate the comments. We will add the reference: "Smith, 1994".

Pg. 6 line 6. *"…. will raise the level of uncertainty in the results"* → *In this case, I would say the reliability/ accuracy of results more than uncertainty.*

- Corrected

Pg. 6 line 11 *"...such as the ANUFLOOD methodology, the Geoscience Australia model, and NSW government curves"* → I understand that no specific literature on models derivation is available but I guess that at least they are cited in some documents, please supply references.

- We totally agree with the reviewer's comments about the ANUFLOOD method. However, the other methods are from some governmental reports that represent only some damage curves without any explanations. In different projects, Australian end-users derive the models or refer to them without having access to the logic behind the model. Therefore, the sentence will be revised and references will be added as "...such as the Geoscience Australia model (Geoscience Australia, 2012), and NSW government curves(Office of Environment and Heritage; New South Wales Government, 2007)".

Pg. 6 line 16 *"Although the detailed valuation survey proposed by Smith seems a little complicated and time-consuming even for data gathered from one type of building (Merz et al.,2010), the new model for evaluating the assembly items and tracking the vertical parameters by considering more*

general categories, has attempted to simplify the process as much as possible→ repetition, see pg. 5 line 11

- We appreciate the reviewer's comment. This sentence will be deleted from page 5: "However, the authors have improved this model and compared it to existing Australian approaches."

Section 3

I think that this section must be changed. See above.

- Corrected. Please see above.

Pg. 6 line 23 → comprehensive or generic????

- Corrected. "Comprehensive" has been deleted.

Pg. 6 line 23 "*Some comprehensive generic depth–damage curves for south-east Queensland have been presented in the report by Geoscience Australia"* → which report?

- Reference will be presented. It has been also cited in page 7 line 3 "(Geoscience Australia, 2012)"

Pg. 7 line 7 *"they are adapted to our area of study"* → By who? What does it means?

- The sentence will be revised as: "they are prepared by the Geoscience Australia for use in our area of study ".

Pg. 7 line 8 "*they are supported by the Australian government"* → And so what?

- This item will be deleted in the next version.

Section 3.2

It is not clear to me how the four classes were selected if you avoided the survey.

- We appreciate the comment. We will add the following reference and explanations: "This selection has been done based on the data collected from the national exposure information system of Australia. This dataset shows that 74 percent of residential buildings in our area of study are made with masonry and timber walls. Also, 99 percent of them are one and two storey buildings (97% one storey and 2 % two storey)."

(Dunford, M.A., Power, L. & Cook, B., 2014. National Exposure Information System (NEXIS) Building Exposure - Statistical Area Level 1 (SA1).Geoscience Australia, Canberra. http://dx.doi.org/10.4225/25/5420C7F537B15)

"As discussed further below, this categorisation is totally adaptable with the empirical flood loss data. In other words, the empirical datasets have expressed the percentage of damage and the condition of flooded buildings based on the damaged sub-assembly groups" → not clear, please specify

- As explained in Section 4.2, the extent of structural damage in the collected datasets has been expressed by some descriptive terms such as: undamaged, minor, moderate, severe, and total damaged. Also, an attached guideline explains these terms based on the affected components.

Considered sub-assemblies (e.g. foundation, below first floor, structure, interiors or exterior walls) in the guideline of empirical datasets are the same as general assembly groups that we have defined.

"In the above formula, the maximum value of damage for each class of building is extracted from the Geoscience Australia report (Geoscience Australia, 2012), which represents the value of damage corresponding to the maximum depth of water (maximum height of the building relative to the first floor)"
→*In order to be coherent with the damage estimation approach, Dmax should be equal to the total value of assembly items. Why this choice?*

- As mentioned by the reviewer, D_{max} would be equal to the total value of damageable assembly items. In other words, d_h is the percentage of damage corresponding to the depth of water (h). Accordingly, D_{max} should be the percentage of maximum damage corresponding to the maximum height of the building (H). On the whole, this formula, by selecting an appropriate value for the root of function "r", tries to find a simple methodology for distributing the maximum percentage of structural damage to the height of the buildings (it would be a function of the damageable property of structures).

" By following the notion that in uniform residential buildings with more than one storey, the first floor of the building contributes more damage than the other stories because most utilities are stored there, this formula enables the user to define how much damage would occur between the first floor elevation, and how much damage can be distributed among the other floors"
→ what are uniform buildings?

- It addresses typically urban buildings that are generally uniform from the second floor. We will provide additional explanation.

→ The concept is not clear. The possibility recalled in the paper is offered by the division of the total maximum damage in the maximum damage at each floor. Please clarify.

- Usually the first floor contributes more damage than other storeys because most electrical equipment is stored here as well as in the basement. Therefore, this formula enables the user to define how much damage would occur between the first floor elevation and the top of the rafters of the first floor, and how much damage will distribute typically among the other storeys.

Section 4

Pg 13 line 6 → *"In addition to some issues regarding the standards of insurance companies that effect their methods of data gathering and collection"* → not clear, please comment on

- We are very grateful for this comment.We will provide additional explanation as: " Insurance companies mainly focus on the collection of data on monetary losses (repair costs) and their relation to the total insured value of the damaged object; while datasets that were gathered with the aim of classifying structural damage or deriving loss estimation models also contain information about the flood characteristics, building types, construction materials, etc. (Thieken et al., 2009)."

Pg. 13 line 13 *"After discarding the unrelated cases"* → what do you mean with unrelated cases?

- Data related to buildings with different uses (e.g. commercial or industrial) or characteristics (e.g. materials and number of storeys). We will provide additional explanation.

Pg, 13 line 17 *"It is to be noted that for selecting the most probable datasets, empirical samples with very rare population have been omitted from this group"* → not clear, what do you mean with "most probable datasets" and "empirical sample with very rare population"?

- Thank you for your valuable comment. As far as you know, each damage function has uncertainties in regard of the data used for the calibration. Due to this matter, presenting damage functions as "minimum", "maximum" and "most-likely" is a selected common practice for some organisations such as the U.S Army Corps of Engineers. However, population of data that show the minimum and the maximum trend was very rare in our dataset (*less than 3%*). This matter could be related to the very low variation of buildings in our areas of study in terms of "building type"; "structural value"; "building quality"; "building size"; and "building age" (Dunford et al., 2014). Consequently, the minimum and maximum curves could not be derived from this rare population, but it can be claimed that the newly derived model has been calibrated with the most likely trend. Also, this rare population (3%), even after resampling of dataset by means of bootstrapping, did not change the considered trend. Following these explanations and for preparing the most likely trend, we have not considered them. In order to avoid confusion in the reader, we will provide these explanations in the next version as well.

Pg. 14 line 10 " Table *1 summarises the contribution of sub-assembly replacement values as a percentage of the total building replacement value"* → I would expect that the contribution of sub-assembly items changes in the four vulnerability class . Did you consider this?

- The reviewer is correct and we have considered this matter. Table 1 has summarised these percentages, on average.

Pg. 14 line 11 *"Accordingly, based on the total value of affected items compared to the entire value of the building (Jonkman et al., 2008), each condition rate has been linked to one range of damage percentage and water depth (see Fig. 3)"* → this point must be clarify. I do not understand the figure. Maybe supplying some numerical examples can be useful.

- Reviewer is correct. For resolving this confusion, we will add more explanations in the next version. On the whole, for exchanging the qualitative terms of losses into percentage of damages, sub-assembly approach proposed by the chapter 4 of HAZUS technical manual has been followed. In this regard, the following steps have been accomplished:

1. For every vulnerability classes, the replacement value of each sub-assembly group compared to the total value of the building has been estimated;
2. For each damaged building, attached guideline indicates which groups of sub-assemblies (e.g. foundation, below first floor, structure, interiors or exterior walls) are partially or totally damaged. For instance, the guideline mentions that for a building with the "moderate" damage condition, following issues will be occurred: electrical and mechanical equipment are impacted (*High set: switchboard inundated; Low set: power points inundated, power supply has been cut*); internal fixtures are damaged; interiors and exteriors have started damaging from the first few centimetres of water depth (*the "minor" category*) but are still not totally damaged (*interiors will be totally damaged in the "severe" category and exteriors will be totally damaged in the "total" category*) ; the structure and foundation are fine but the building is not habitable and occupants may need to vacate during repairs

3. Based on the guideline description and the relative values of affected items, condition rate has been linked to the percentage of building loss.
4. This calculation will be repeated for all damage categories and based on that, the Figure 3 can be depicted for all vulnerability classes. In this figure, the horizontal axis is the overall building loss as a percentage of building replacement value and the vertical axis is the sub-assembly loss as a percentage of its own replacement value.
5. For every building, on the basis of the reported water depth and the recorded damage condition, the percentage of damage vs. depth of water could be extracted.

Pg. 14 line 14 *"Finally, for every building, based on the magnitude of hazard and depth of water, the percentage of damage could be extracted"* → not clear. Please specify

- Please see above.

Section 5

Pg. 15 line 6. *"As stated earlier, the first step of model derivation is choosing the maximum possible value of damage that can occur for each vulnerability class"* → where?

- We appreciate the comment. "As stated earlier" will be deleted from this sentence.

Pg. 15 line 10 *"Due to the fact that the utilities of buildings (including mechanical and electrical facilities) are mostly stored on the first floor of buildings, and based on the replacement value of this equipment compared to other fragile items (e.g. the superstructure, exteriors and interiors excluding utilities), the magnitude of first floor damage relative to second floor damage can be estimated to be approximately 1.8 times more"*. → I expect this to change according to the material. Did you consider this?

- The reviewer is correct. This ratio would vary according to the material of buildings and we have considered it.

Section 6

I guess two different sets of data have been used to calibrate (i.e. defining the r values) and validate (i.e. comparing with empirical data) the model but this is not clear in the paper. Otherwise, it is evident that the new model better predict expected damage. It can also be useful to know which is the dimension of the datasets used for calibration and validation, for each class.

- Reviewer is totally correct and we have substantially revised this part. Please see the above explanations.

"Due to the fact that potential damage is the maximum possible value of losses without considering any mitigation measures (Bureau of Transport Economics, 2001; Molinari, 2011; Molinari et al., 2013), underestimated values which represent the percentage of damages less than actual values should be omitted for the averaging part" → why? Not clear to me, this way you are not assessing the real performance of the model

- We appreciate your comment. Based on the above explanations and your suggestion, section 6 and the related comparisons have been revised substantially.

"As can be seen from Table 3, the average values have been calculated from the ratios greater than one, and the standard deviation and coefficient of variation have also been estimated based on these ratios" → the table is not clear as all the ratios are reported. Moreover, results should be checked. For example, the average of positive ratios for model GA is 2.68 not 2.58

- We are very grateful for this comment. For resolving this confusion, results representation will be revised and it will be presented in a better manner.

"Figs. 7 and 8, all approaches overestimate the magnitude of losses for the first few centimetres of flood (approximately the first 15 cm)" → this is true for Figure 8 but not for Figure 7 where the three models overestimate in the first 75 centimetres. A model that cannot be used in the first 15 cm may have sense where expected water levels are significant (as in the case of Australia). But neglecting the first 75 cm means that the model cannot be used in many circumstances. What is the authors intent in these cases?

- We appreciate the comment. Based on the new dataset provided for the model calibration and with the aim of resolving this issue, values of root functions have been modified.

"the newly derived model represents lower values for the average of the loss ratios and they are closer to one (i.e. the loss ratio is equal to one if the results of the functions and empirical datasets match each other)" → if you are not using different dataset for calibration and validation, of course it does. The other models were not calibrated with empirical data!!! See above

- We have considered this comment. Please see the above explanations

"we see that the newly derived model is also conservative, and marginally less than the results for the GA method and the USACE approach" → what do you mean with conservative?

- Conservative means overestimating the magnitude of potential damage. We will revise this sentence as: " the newly derived model also overestimates the magnitude of damages, but marginally less than the results of the GA method and the USACE approach"

"Other methodologies that represent unstable trends seem to be more uncertain and difficult for making decisions" → not clear, please specify

- By the unstable trend, we mean that for some intervals of water depth, the damage model considerably underestimates the magnitude of losses; while for some intervals, it significantly overestimates the extent of damages. Therefore, this fluctuation can bring uncertainty to the results and it will affect the quality of decisions taken for risk reduction purposes. In order to avoid confusion in the reader, we will provide more explanations in the next version.

<u>Conclusion</u>

"This study aimed to investigate the applicability and adaptability of different flood damage models to Australian geographical conditions" → *adaptability is not analysed in the paper (see above)*

- Objective of study has been modified based on your suggestion.

"The results of this study show that even the state methodologies will express the results of flood damage conservatively, either underestimating values or sometimes producing marginally high values" → what do you mean with conservative (see also above)? Always overestimating? Or always underestimating? The two features of the model described in the sentence are in conflict.

- We are very grateful for this comment. This confusion will be resolved by modifying the sentence as: "The results of this study show that even the state methodologies might considerably overestimate the magnitude of flood impacts, or significantly underestimate the value of losses, if they have not been adapted with empirical datasets."

Bibliography

I did not check the bibliography at this stage of the review. I reserve to do this in a second time. However I noted that no reference is made for most of the discussed tools (i.e. damage models).

- We will provide additional references.

Figure 6

It is very strange to me to see empirical data which are so well aligned (i.e. without spread). Probably they have been filtered in some way but this is not clear in the paper.

- We appreciate the comment. These curves, due to the above issues, have been changed substantially.