Response to the reviews

We would like to thank the Referees for their careful and interesting evaluation of this Brief Communication; their remarks and advice will help us to make the paper more clear, readable, and complete. In the following, a point by point answer to questions and comments raised by the referees is supplied.

The numbers of lines outside the brackets refer to the marked-up version of the manuscript, in brackets to the new version.

Referee #1

R1-C1: I would suggest to add a few sentences in the discussion section on the general added value of the paper (beyond the exclusive applicability in Italy), namely that complexity reduction or the development of surrogate models can be recommended without losing too much accuracy. This is especially the case, if not all of the independent variables are available as model input.

Thank you for the suggestion, we agree that highlighting the strength of the work, also as a method to develop surrogate models without losing too much accuracy, could increase its value. We added a sentence in the discussion section to underline this concept, lines 159-160 (128-129).

R1-C2: Another suggestion is to add a reference to the dataset you mentioned in lines 87-88, the "real estate data of Northern Italy".

Thank you for the comment. In fact, "real estate data of Northern Italy" could be misleading, as the reader could think that there is a unique database with all the information. But the types of probability distributions and their parameters were chosen on the basis of information coming from different databases. For instance, the percentage of buildings with high or low level of maintenance was computed according to the provincial data of Istat (Italian institute of statistics) for Piemonte, Lombardia, Emilia Romagna and Veneto region; for the estimation of the percentage of building with high or low finishing level, we associated the finishing level to the cadastral categories supplies by the real estate market observatory of revenue agency (Osservatorio del Mercato Immobiliare OMI by Agenzia delle Entrate) in the same regions. In the paper, we will be more precise about the source of information, lines 92-95 (85-88). However, if the method wants to be repeated, different probability distributions and information sources can be used, according to the territorial context and available databases.

Referee #2

R2-C1: The sensitivity analysis is not very clear. What is the basis for choosing the predictors? Also, does the model performance becomes worse when one/some of the chosen variables are treated as default values? I think, a clearer explanation will help to extend this approach to other synthetic models. Not many synthetic models allow missing variables. So, some details on how to choose the default values may be helpful.

We chose the predictors and the implicit variables according to different criteria, explained in lines 59-71 (55-67). For instance, we neglected variables that are not available or considered fixed or that are functions of other variables. We implemented a local sensitivity analysis consisting in perturbing one single parameter around its reference value, keeping the others constant, to evaluate the resulting variation of the output. The parameters that do not cause significant changes in damage estimation were neglected. In particular, the variables that were finally included in simple-INSYDE cause a variation of damage between 12% and 38%, while neglected ones less than 2%. In the original version of the paper, we did not describe the sensitivity analysis in detail because of the brief and concise nature of the manuscript, but we understand that is too synthetic; in the new version we added some information without going into details, lines 77-79 (71-74). Still, it is worth noting that the implemented sensitivity analysis is one possible method to select relevant variables; other modellers may choose other procedures they consider most appropriate.

Regarding the model performance, we did not evaluate it in the case also the chosen explicative variables are set as implicit (e.g. in the case the model became bi or uni-variable), but the simple structure of the model allows to easily implement this type of analysis, lines 131-132 (113-114).

R2-C2: Transferability appears to be an advantage of INSYDE (Dottori et al. 2016, Amadio et al., 2019). I agree with the authors that there is a need to reduce complexity in order to make the model widely applicable. But, feeding in more of local inputs as default values may cause serious bias when the model is transferred, as such. The authors mention about this. But, an illustration of how to work around this limitation will add value to the study.
As explained in the discussion chapter, we think that the simplification of INSYDE could make the model widely applicable in the context of calibration, Northern Italy, but it does not make the model more transferable in other contexts. To transfer the model, it is necessary to adapt the original INSYDE to the context of interest, as mentioned in Dottori et al. (2016), and then to apply the procedure shown in the paper with assumptions coherent with the development context.

R2-C3: It is also not clear how the calibration of simple-INSYDE is implemented (line: 84). Is the INSYDE original model applied with the same set of variables (Table 2) considering the rest as missing/default and then an interpolation is performed or are there additional variables involved?
Yes, the variables that are not in Table 2 are considered as default. The model INSYDE and simple-INSYDE work with the same set of buildings and flood features. The values of the functions' coefficients were manually changed to reduce the error. We modified the part of calibration in order to make clearer the procedure, lines 89-92 (82-85).

R2-C4: Since both the INSYDE model development and the simplification approach are based on same regions (Northern Italy), it is difficult to judge the general applicability of this simplification approach based on the reported errors. The comparison with same set of simulated buildings appears like evaluating the fit of the interpolation function with same train and test data. An alternative is that the authors may consider providing validation on real damage data like Amadio et al (2019).
Thank you for the comment. The aim of the comparison was to evaluate the error if we use simple-INSYDE instead of INSYDE, for this reason we used the same set of building parameters and flood parameters. We assumed that the performance of the original INSYDE is good in the panorama of Italian flood damage models, without providing an evaluation as it has been already done in other papers quoted in the introduction: e.g. Amadio et al., 2019, Molinari and Scorzini, 2017, Molinari et al., 2020.

R2-C5: Since INSYDE is a probabilistic model, some discussions on uncertainty in predictions from both INSYDE and simple-INSYDE will be an interesting addition to the discussions.
INSYDE is a probabilistic model because some damage mechanisms are modelled using probabilistic functions rather than deterministic (Dottori et al., 2016). Simple-INSYDE was developed through deterministic interpolations of the model functions. The investigation of how uncertainty in damage predictions propagates into the output is interesting but out of the scope of the paper.

R2-C6: 1. Some insights into 'why' INSYDE is complex and hard to implement (what aspects?) may add value to the objective.
Thank you for the suggestion. The main feature behind INSYDE complexity is its articulate structure, made of a lot of functions. This fragmented structure prevents to understand the relation between damage and its explanatory variables. Moreover, the model is now available in R. Simpler functions could be implemented and understood also by users that do not know R or that want to integrate the damage functions with other analysis tools. We added a sentence to clarify this aspect, lines 20-21 (19-21).

R2-C7: 2. Line 16: reparation may be replaced with repair
Yes, thank you for the advice.

R2-C8: 3. Line 21: How does this improvement help integration to a GIS software? This is not discussed anywhere.

The reference to GIS software was an example, but we see that it is not clear. The idea is that simpler functions are simpler to be implemented with other programming language or to be integrated with other tools, as the GIS tool. We changed this sentence, lines 20-21 (19-21).

R2-C9: 4. Lines 21-24: Reduction in dimensionality of the INSYDE model should be included. The model doesn't completely preserve the multi-variable nature of INSYDE. Some variables are treated as constants in the simple model.

Yes, simple-INSYDE considers less variables than INSYDE, but its nature remains multi-variable because it considers several damage explanatory variables and represents damage mechanisms that are functions of a lot of variables, even if some parameters are implicitly considered.

R2-C10: 5. In Table 1, it is difficult to understand X. From the context, I understood that these are values user has to input. A note will help.

Thank you, "x" is not clear, we added a note with x meaning in Table 1.

R2-C11: 6. Line 45: walls and plants? I think it is a typo.

Thank you, we replaced plants with systems (we mean electrical or heating system).

R2-C12: 7. Line 51: Footprint Area is interchangeably used as FA and A (table 1 and 2)

Yes, we used A and not FA for the simplified model. We added a note in Table 1.

R2-C13: 8. Line 70: There is no quantification provided for sensitivity analysis. Hence, the context for 'significantly' is missing.

As said in R2-C1, we did not describe in detail the sensitivity analysis because of the short nature of the paper and because it is a possible way to select the variables, but not the unique. However, we enriched the description, also providing some quantitative performance indexes, lines 77-79 (71-74).

R2-C14: 9. Lines 91-94: The range of acceptable errors is very huge. Given this argument, even the need for important variables considered in simple-INSYDE may be questioned

Yes, the range is huge, but it refers to an example from the application of the FLEMO-ps model (Thieken et al., 2008); the example was chosen to emphasize that uncertainty in flood damage estimation could be very high. The discussion about the choice of considering very few or several variables, so to discuss advantages and disadvantages of uni- or multi-variable models, is interesting but it needs further analysis and discussion that transcends the objective of this paper. According to the choice of a brief communication, we prefer focusing on the main objective of the study.

R2-C15: 10. Table 2 needs reference. Also, please introduce a column with full-forms to make it easy for the reader.

There is not a specific reference, we choose the value of probability distributions according to data about building characteristics and real estate market in some regions of Northern Italy (see R1-C2). We changed the sentence about the calibration and Table 2 in order to be more precise about the data sources and the probability distributions, but we prefer to use the synthetic notation to reduce space and because the full-forms are easily available in statistical books or websites.

R2-C16: 11. Lines 100: Please rephrase that simple-INSYDE is a simplified version of INSYDE. The fundamental assumptions and methodologies are from INSYDE.

Ok thank you, we rephrased it.

R2-C17: 12. The arrangement of the Discussion section 3 is not coherent. The model is for Northern Italy. But, more focus on wider applicability of such an approach and how to implement this for other regions will be interesting.

See answer R2-C2.

Referee #3

R3-C1: Introduction: I guess one of the limitation of the original version of the model (INSYDE) might be related to availability of all required data. If this is the case, it could be worth mentioning it in the text. Also, please provide more details regarding difficulties with GIS application.

About the obstacle linked to the availability of the required data, the original INSYDE overcomes this by setting default values for missing variables. About the GIS application, it was an example, but we see that is not clear (see also answer R2-C8). The point is that simpler functions are simpler to be implemented with other programming language or to be integrated with other tools, as the GIS tool. We changed this sentence to clarify the concept, lines 20-21 (19-21).

R3-C2: Introduction, L19-21: the presentation of the limitations that inspired the new version is quite general. Can you better specify them?

The main reason that inspired the new version is the articulated structure of the INSYDE model. We think that a simpler structure could facilitate the analysis of the role of the variables in damage computation and the easily implementation of the model. See also R2-C6.

R3-C3: Table 1: table caption should also include the simple-INSYDE. Also, referring to simple-INSYDE, it might be helpful distinguishing independent hazardous variables respect to variable representative of the exposure. Please, specify the meaning of x used in table 1.

Yes, thank you for the advice, we adjusted the table and added the meaning of x.

R3-C4: Some of the variables assumed with a fix value in simple-INSYDE were constant also in the original model (see e.g. IH, BH, GL). What is the difference compare to INSYDE? What are the fixed values adopted for simple-INSYDE?

In INSYDE, the user can choose the value of the variables or use the default values, and this is valid also for variables as IH, BH, GL. In simple-INSYDE, they are implicitly considered so it is not possible to change them. The fixed value adopted in simple-INSYDE are the default values of INSYDE.

R3-C5: Apart from a contained complexity, are there advantages on using a fixed values for same variables that can be directly estimated from other required in any case by the simple version? I am thinking for example to IA or EP, which depend on FA that is still necessary for the application of simple-INSYDE. Also in this case, how did you define those default values? Are they based on observations or assumptions?

We apologize if the answer will not be thorough, but we did not fully understand the request of the referee. Variables as IA or EP or IP refer to geometric features of the building that, if unknown, INSYDE computes by means of some functions of FA. To develop simple-INSYDE we did not define these values, but we decided to maintain the original functions of FA, to be coherent with the original model, lines 70-71(66-67).

R3-C6: L51: footprint area in table 1 is indicated as FA, not A. Also check Table 2.

Yes, we used A and not FA for the simplified model, but we see it is confusing. We added a note in Table 1. See also R2-C12.

R3-C7: L53: wouldn't be more correct saying "flood affected storeys"?

We used the term "exposed to flood" and not affected, because we can use the damage model for damage forecast, and the storeys are not yet affected, but exposed to the event.

R3-C8: L55: wouldn't be more precise saying "independent hazardous variables"?

We mean all the variables, that refer both to hazard and vulnerability.

R3-C9: L62: I suggest providing more details on these values. This would help the possibility to, eventually, apply the model elsewhere.

The variable which were not selected as independent variable were set at default values defined in INSYDE, to be coherent with the original model. We added a sentence to clarify the choices of the independent variables, lines 62-63 (58-59). These choices are strongly related to data availability in the region of interest and on assumptions about the typical building typologies, but, in another region and with another model, the criteria could be different. See also R2-C1.

R3-C10: L73-76: honestly, these steps are not clear to me. If you can extend the explanation, or provide an example, it would be much easier to understand the procedure.
We enriched the explanation but without getting too detailed in the description of the procedure for the definition of each function, in order to be coherent with the choice of a brief-communication, lines 82-89(76-82).

R3-C11: Eq. 3): what happen in case LM is medium?
LM medium is not considered in simple-INSYDE. The user must choose between low or high.

R3-C12: Table 2: in order to make the model transferability easier it would be worth reporting in table 2 the range of the values considered for all the variables.
For the variables that are not present in Table 2, the values are set as default, we added a sentence to clarify this, lines 92(85). See also R2-C3.

List of all relevant changes made in the manuscript

In the following, the list of the relevant changes made in the manuscript is supplied. The numbers of lines outside the brackets refer to the marked-up version of the manuscript, in brackets to the new version.

- Lines 20-21 (19-21): In the Introduction, we added a sentence to enrich the explanation of the limitations due to the complex structure of the original model.
- Lines 62-63 (58-59): We added a sentence to clarify the criteria of selection of independent variables in the simplified model version.
- Lines 77-79 (71-74): We enriched the explanation of the sensitivity analysis, without going into details to maintain the brief nature of the manuscript.
- Lines 83-89 (77-82): To make clearer the procedure of definition of the interpolating functions, we rephrased the explanation and added some examples.
- Lines 89-95 (82-88): We moved the lines about the calibration procedure before the equations and we added a sentence about the data sources of the probability distributions of the damage explicative variables.
- Table 2: We modified Table 2 to make it clearer using a synthetic notation for the probability distributions.
- Lines 121-128 (101-108): We moved some lines about calibration results after Table 2.
- Lines 155-171 (123-139): We moved part of the discussion after Figure 1.
- Lines 159-160 (127-128): We added a sentence in the discussion to underline the concept that the simplified model version does not significantly decrease the accuracy of the model.
- Lines 170-171 (138-139): We added a sentence to highlight the possibility to develop the proposed methodology for other multi-variable models as well.

Below, the marked-up version of the manuscript.

Brief communication: simple-INSYDE, development of a new tool for flood damage evaluation from an existing synthetic model

Marta Galliani, Daniela Molinari, and Francesco Ballio

Department of Civil and Environmental Engineering, Politecnico di Milano, Milan, 20133, Italy

Correspondence to: marta.galliani@polimi.it

Abstract

INSYDE is a multi-variable, synthetic model for flood damage assessment to dwellings. The analysis and use of this model highlighted some weaknesses, linked to its complexity, that can undermine its usability and correct implementation. This study proposes a simplified version of INSYDE which maintains its multi-variable and synthetic nature, but has simpler mathematical formulations permitting an easier use and a direct analysis of the relation between damage and its explanatory variables.

1 Introduction

INSYDE (IN-depth SYnthetic Model for Flood Damage Estimation, Dottori et al., 2016) is a synthetic model for the estimation of flood damage to residential buildings at the micro scale (i.e. building level), developed and tested in Italian case studies (Amadio et al., 2019, Molinari and Scorzini, 2017, Molinari et al., 2020). The monetary damage to a dwelling is computed in the model as the sum of 33 different components, referring to the costs of -repair~~reparation~~, removal and replacement of the damaged elements, which are functions of several damage explicative variables, related both to the hazard and to the vulnerability of the affected item (Table 1). Since the same explicative variable may directly or indirectly influence more than one damage component, it is difficult to understand the weight that each explicative variable has on the overall damage estimate. Moreover, the complex and articulate structure of INSYDE could dissuade less expert users to use the model and discourage its implementation with platform and calculation tools others from the original one. ~~Moreover, the complex and articulate structure of INSYDE could discourage the implementation of the model and its use through other platform such as GIS software.~~ This study proposes an alternative version of the model, named simple-INSYDE, which aims at overcoming these limitations. Simple-INSYDE preserves the multi-variable nature of the model, but aggregates damage components in a smaller set of functions, which clearly describe the role of each explicative variable on the total damage figure and can be easily implemented, even by non-expert users. Such functions are calibrated for low-velocity floods, with building characteristics typical of Northern Italy. The method and the assumptions implemented to obtain the simplified version of the model are described in the following sections.

Table 1. Input variables considered in INSYDE (Dottori et al., 2016), and selected variables in simple-INSYDE. “x” refers to the variables chosen in simple-INSYDE as independent variables.

Event features and building characteristics variables in INSYDE		Default values in INSYDE	Independent variables chosen in simple-INSYDE
h_e	water depth outside the building [m]	[0;5] Incremental step: 0.01m	x
h	water depth inside the building [m]	h = f (h_e, GL)	Fixed at default value
v	maximum velocity of the water perpendicular to the building [ms^{-1}]	0.5	Fixed at default value
s	sediment load [% on the water volume]	0.05	Fixed at default value
du	duration of the flood event [h]	24	x
q	water quality (presence of pollutants)	Yes	x
FA	Footprint Area [m^2]	100	x (renamed A in simple-INSDYE)
IA	Internal Area [m^2]	0.9 FA	Fixed at default value
BA	Basement Area [m^2]	0.5 FA	x
EP	External Perimeter [m]	4√FA	Fixed at default value
IP	Internal Perimeter [m]	2.5 EP	Fixed at default value
BP	Basement Perimeter [m]	4√BA	Fixed at default value
NF	Number of floors	2	Functions for one storey
IH	Interfloor height [m]	3.5	Fixed at default value
BH	Basement height [m]	3.2	Fixed at default value
GL	Ground floor level [m]	0.1	Fixed at default value
BL	Basement level [m]	– GL – BH – 0.3	Fixed at default value
BT	Building type (1 detached house, 2 semi-detached, 3 apartment)	1	Fixed at default value
BS	Building structure (1 reinforced concrete, 2 masonry)	2	x
FL	Finishing level (1.2 high, 1 medium, 0.8 low)	1.2	x
LM	Level of maintenance (1.1 high, 1 medium, 0.9 low)	1	x
YY	Year of construction	1994	Fixed at default value
PD	Heating system distribution (1 centralized, 2 distributed)	1 if YY≤1990, 2 otherwise	Fixed as 1 (centralized)
PT	Heating system type (1 radiator, 2 pavement)	2 if YY>2000 and FL>1, 1 otherwise	Fixed as 1 (radiator)

2 Method

The first step to provide a simpler structure of the model was to aggregate the original damage functions into four components:

- Damage to basement: in case of flood, basement is assumed totally inundated and damage does not depend on water level.
- Damage to floor: in case of water level higher than the level of floor, the damage to floor is counted as independent from water level.
- Damage to storey: it considers damage to the elements over the floor (e.g., walls and systems ~~plants~~) that depends on water level.
- Damage to boiler: it depends on water level only if the basement is not present, otherwise, the boiler is considered located in the basement which is completely inundated.

In order to support model transferability (Merz et al., 2010), the simplified model computes damage in relative terms, as the ratio of the absolute damage to a reference value. The reference value is set as the cost of reconstruction of the storeys exposed to the flood, which~~; cost of reconstruction~~ is evaluated as the product of the replacement value RV [€/m^2] and the footprint area A of each storey [m^2]. Equation (1) represents the conceptual formula of the simplified model, where D is the building damage in absolute term [€], d in relative term, n is the number of flood exposed storeys.

$$D = RV_{basem} \cdot A_{basem} \cdot d_{basem} + RV_{storey} \cdot A \cdot \left(\sum_{i=1}^{n} \left(d_{storey_i} + d_{floor_i} \right) + d_{boiler} \right) \tag{1}$$

The second step was the choice of the independent variables to be included in the model, among those of the original INSYDE (Table 1). The variables that were not included in simple-INSYDE were not effectively neglected, but implicitly assumed at the default values according to the assumptions made in INSYDE, for the geographical context and the flood type of interest (Wagenaar et al., 2016). We fixed as default value the variables considered constant in the geographical context, or that are generally not known, or that are functions of other variables or that do not influence significantly damage simulation. Among the ~~event feature~~ variables describing the event, we preserved the water level, the duration of the flood and the presence of pollutants. Indeed, the sensitivity analysis performed in Dottori et al. (2016) highlighted that, in case of slow riverine flood events, water velocity and sediment load have a minor influence on damage, compared to the chosen variables. The selection of the vulnerability variables followed different criteria. We considered the interfloor height and the basement height fixed at their default values because they do not vary significantly in Northern Italy. We kept the default value also for the ground floor level and the heating system variables (PD and PT), because information on them is difficult to retrieve, without a detailed field survey. The internal area, the external perimeter, the internal perimeter, the basement perimeter and the basement level are fixed as functions of other variables in INSYDE, and we maintained the same functions ~~this assumption~~. However, this implies limiting the use of the model for the estimation of damage to single housing units, and not to condominiums; indeed, the functions to estimate perimeters in INSYDE were established considering the typical configuration of a 100 m^2 detached Italian house; this configuration is kept constant in the model, thus not considering a variation in the number of rooms or a multiplication of housing units in case the building area increases. ~~The remaining vulnerability variables were the object of a sensitivity analysis, which revealed that the year of construction and the building type do not significantly affect damage estimate.~~ The remaining vulnerability variables were the object of a one-at-a-time sensitivity analysis, which revealed that the variation of the values of the year of construction and of the building type causes a change of the mean relative damage smaller than 2%, compared to the caused by the other variables which is between 12% and 38%. On the other hand, the building type, in Italy, is important to evaluate the replacement value. Table 1 shows the variables that were finally considered in simple-INSYDE.

The last step was the development of the simplified functions. For the four damage components, a set of reference values was defined for each variable that influences the component under investigation, e.g. A=100 m^2 and FL=low. Then, damage to each component was calculated ~~component was computed~~ by varying the value of one variable at a time (e.g. f(A) or f(FL)) and keeping the others at their reference value. ~~its input variables one by one around the reference value, in order to identify~~ The resulting functions are simple interpolating functions suitable for representing the role of each variable in the estimation of a specific damage component: for ~~on the final damage figure~~ variables that assume only two values, as BS or FL, the only identification of multiplicative coefficients was required, for variables with a wide range of values, as d or A, the functions that approximate the role of the variables in the final damage figure are more complex. Then, ~~T~~the interpolating functions were calibrated comparing the damage simulated by the simplified model and the original model for a sample of 10000 buildings, whose features (i.e. input variables values) were partly randomly selected from probability distributions assumed representative of Northern Italy (Table 2) and partly ~~.~~ set at default values. In particular, the parameters of the probability distributions were chosen on the basis of data supplied by Istat (the Italian institute of statistics), the real estate market observatory of the revenue agency (Agenzia delle Entrate) and some provincial databases that refer to built environment in the regions Emilia Romagna, Lombardia, Piemonte and Veneto. ~~Results~~

The final result of the process is ~~are~~ expressed by equations 2-5, which represent the simple-INSYDE model:

$$d_{basement} = f(A_{basem}) f(du) \rightarrow \begin{cases} f(A_{basem}) = 0.02 + \frac{0.35}{\sqrt{A_{basem}}} \\ f(du) = 1 + 0.3 \arctan(du - 36) \end{cases} \quad (2)$$

$$d_{storey} = f(h)f(A)f(LM,du)f(BS)f(FL)f(q) \rightarrow \begin{cases} f(h) = \left(0.17h - 0.02h^2\right) \\ f(A) = \left(0.2 + \frac{7}{\sqrt{A}}\right) \\ f(LM,du) = \begin{cases} 1 + 0.15 \cdot \arctan(du - 36) & \text{if LM low} \\ 0.8 + 0.2 \cdot \arctan(du - 36) & \text{if LM high} \end{cases} \\ f(BS) = \begin{cases} 1.35, & \text{if BS masonry} \\ 1, & \text{elsewhere} \end{cases} \\ f(FL) = \begin{cases} 1.5, & \text{if FL high} \\ 1, & \text{elsewhere} \end{cases} \\ f(q) = \begin{cases} 1.2, & \text{if } q = 1, \text{presence of pollutants} \\ 1, & \text{elsewhere} \end{cases} \end{cases} \quad (3)$$

$$d_{floor} = f(h, FL) = \begin{cases} 0.04, & \text{if h>0 and FL high} \\ 0, & \text{elsewhere} \end{cases} \quad (4)$$

$$d_{boiler} = f(A_{basem}, h) = \begin{cases} 0.015, & \text{if } A_{basem} \neq 0 \text{ or } A_{basem} = 0 \text{ and } h > 1.6 \text{ m} \\ 0, & \text{elsewhere} \end{cases} \quad (5)$$

where the units of measures of the variables are m^2 for the area (A), hours for duration (du), and m for water depth (h).

~~The interpolating functions were calibrated comparing the damage simulated by the simplified model and the original model, for a sample of 10000 buildings, whose features (i.e. input variables values) were randomly selected from probability distributions assumed representative of Northern Italy (Table 2). In particular, for the footprint area, the finishing level, the building structure and the maintenance level, the distribution parameters were chosen on the bases of real estate data of Northern Italy. The comparison of the simulated damage by the original and the simplified model, showed a mean relative error equal to 0.24, with a ratio to the mean absolute damage equal to 1.07. The application of the model INSYDE in real case studies (Dottori et al., 2016, Molinari and Scorzini, 2017, Amadio et al., 2019, Molinari et al., 2020) showed good performance of the model, with a mean ratio between the total damage simulated and the observed damage equal to 1.26. On the other hand, literature shows that the performance of flood damage models can be affected by high uncertainty, with relative errors vary from 20% to exceed 1000% (Scorzini and Frank, 2017, Thieken et al., 2008). Thus, we consider that the additional error caused by the use of simple-INSYDE is acceptable, and that the estimation of the overall damage is comparable with that supplied by INSYDE.~~

Table 2. Probability distributions and respective parameters of the explicative variables.

~~Variable~~	~~Distribution~~	~~Parameters~~	~~Notes~~
~~A~~	~~Log-Normal~~	~~μ=5.10, σ^2=0.49~~	~~μ mean, σ^2 variance~~
~~FL~~	~~Bernoulli~~	~~p=0.02~~	~~probability FL high~~
~~BS~~	~~Bernoulli~~	~~p=0.64~~	~~probability BS masonry~~
~~LM~~	~~Bernoulli~~	~~p=0.86~~	~~probability LM high~~
~~NF~~	~~Discrete Uniform~~	~~[1,10]~~	
~~Basement~~	~~Discrete Uniform~~	~~[0,1]~~	~~0 absent, 1 present~~

~~BT~~	~~Discrete Uniform~~	~~[1, 3]~~	~~1 detached house, 2 semi-detached, 3 apartment~~
~~du~~	~~Discrete Uniform~~	~~[10, 60]~~	~~unit of measure: hour~~
~~h~~	~~Continuous Uniform~~	~~[0, 3.5]~~	~~unit of measure: meters~~
~~q~~	~~Discrete Uniform~~	~~[0, 1]~~	~~0 absent, 1 present~~

Variable	Distribution		Notes
A	Log-Normal	Lognormal(5.10, 0.49)	
FL	Bernoulli	B(1, 0.02)	0.02 probability FL high
BS	Bernoulli	B(1, 0.64)	0.64 probability BS masonry
LM	Bernoulli	B(1, 0.86)	0.86 probability LM high
NF	Discrete Uniform	U{1, 10}	
Basement	Discrete Uniform	U{0, 1}	0 absent, 1 present
BT	Discrete Uniform	U{1, 3}	1 detached house, 2 semi-detached, 3 apartment
du	Discrete Uniform	U{10, 60}	unit of measure: hour
h	Continuous Uniform	U(0, 3.5)	unit of measure: meters
q	Discrete Uniform	U{0, 1}	0 absent, 1 present

The comparison of the simulated damage by the original and the simplified model, showed a mean relative error equal to 0.24, with a ratio to the mean absolute damage equal to 1.07. The application of the model INSYDE in real case studies (Dottori et al., 2016, Molinari and Scorzini, 2017, Amadio et al., 2019, Molinari et al., 2020) showed good performance of the model, with a mean ratio between the total damage simulated and the observed damage equal to 1.26. On the other hand, literature shows that the performance of flood damage models can be affected by high uncertainty, with relative errors vary from 20% to exceed 1000% (Scorzini and Frank, 2017, Thieken et al., 2008). Thus, we consider that the additional error caused by the use of simple-INSYDE is acceptable, and that the estimation of the overall damage is comparable with that supplied by INSYDE.

3 Discussion

This study led to the main objective of developing a new tool for flood damage estimation to dwellings from the simplification of the model INSYDE. The new model~~, which~~ is based on a sensible number of available input data and allows investigating the relation between damage and its explanatory variables by means of a simple set of functions. For instance, Figure 1 shows the relative damage computed by simple-INSYDE as a function of water depth, for the different damage components of the model. The figure highlights that the storey component gives the biggest contribution to the damage and it is the only one depending on water level. The other components are independent on water level and have a lower weight on the final damage figure; still, they assume a non-negligible role, especially in case of shallow waters.

Moreover, the study allowed to deeply investigate the behaviour of the original model and to highlight shortcomings that could be further improved in the future. For example, assumptions made in the model on building configuration, which limit its use to single housing units and not condominiums, is not directly reported in the paper of Dottori et al. (2016), but is important for a correct implementation of the model and a better understanding of estimation errors.

Compared to the original model, the simplified model requires fewer input variables, facilitating the model implementation, but impeding the control by the user on the variables that are implicitly considered. For this reason, Simple-INSYDE is less adaptable to contexts different from the calibration one than INSYDE. It is worth recalling that simple-INSYDE is addressed to evaluate damage in case of low-velocity floods and built environments typical of Northern Italy. It is recommended not to use it for other types of inundation (Kreibich and Dimitrova, 2010) or for other types of building and/or geographical contexts. In these cases, the derivation of new interpolating functions from INSYDE, with the process described in this study is suggested; to this aim, the original model needs be adapted to the context of interest, by modifying the default values of the variables and the unit prices of the building components, then, the simplification method can be implemented to obtain new functions with new coefficients.

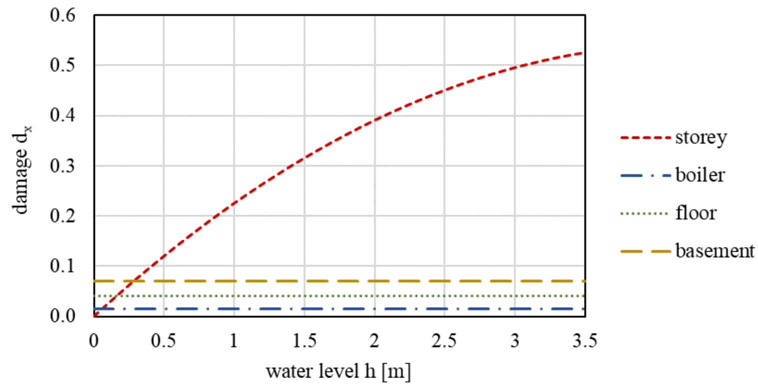

Figure 1 - Comparison of the simple-INSYDE damage components as functions of water depth. Damage values in the figure refer to the case of a 100 m^2 building, with cellar, in reinforced concrete, with high finishing level and low maintenance level, affected by a flood of 36 hours, in absence of pollution.

Moreover, the study allowed to deeply investigate the behaviour of the original model and to highlight shortcomings that could be further improved in the future. For example, assumptions made in the model on building configuration, which limit its use to single housing units and not condominiums, is not directly reported in the paper of Dottori et al. (2016), but it is important for a correct implementation of the model and a better understanding of estimation errors.

The study demonstrates that the use of a simplified model, which is consistent with the assumption of the original one, can lead to comparable results and does not decrease considerably the accuracy of damage forecast.

Compared to the original model, the simplified model requires fewer input variables, facilitating the model implementation, but hamperingimpeding the control by the user on the variables that are implicitly considered. For this reason, simple-INSYDE is less adaptable to contexts different from the calibration one than INSYDE. In this regard, iIt is worth recalling that simple-INSYDE is addressed to evaluate damage in case of low-velocity floods and built environments typical of Northern Italy. It is recommended not to use it for other types of inundation (Kreibich and Dimitrova, 2010) or for other types of building and/or geographical contexts. In these cases, the derivation of new interpolating functions from INSYDE, with the process described in this study is suggested; to this aim, the original model needs be adapted to the context of interest, by modifying the default values of the variables and the unit prices of the building components, and then by implementing, the simplification method can be implemented to obtain new functions with new coefficients.

It is worth noting that the method proposed in this study for the derivation of the model is not limited to INSYDE, but can be repeated and developed also for other multi-variable models to obtain alternative, simpler and more explicit versions.

Acknowledgments

Authors acknowledge with gratitude Anna Rita Scorzini, from the Department of Civil, Environmental and Architectural Engineering, University of L'Aquila, for her fruitful suggestions and hints during the developing of the work.

References

Amadio M., Scorzini A. R., Carisi F., Essenfelder A.H., Domeneghetti A., Mysiak J., and Castellarin A.: Testing empirical and synthetic flood damage models: the case of Italy. Nat. Hazards Earth Syst. Sci., 19, 661–678, doi:10.5194/nhess-19-661-2019, 2019

Cineas-Cresme-Ania.: I costi di costruzione in edilizia residenziale, industriale, per uffici e alberghiera, 2018

Dottori, F., Figueiredo, R., Martina, M. L. V., Molinari, D., and Scorzini, A. R.: INSYDE: A synthetic, probabilistic flood damage model based on explicit cost analysis, Nat. Hazards Earth Syst. Sci., 16, 2577–2591, doi:10.5194/nhess16-2577-2016, 2016.

Kreibich, H., Dimitrova, B.: Assessment of damages caused by different flood types, in Flood Recovery, Innovation and Response II, edited by Wrachien, D., Proverbs, D., Brebbia, C. A., Mambretti, S. (Eds.), WIT Press, pp. 3-11, doi:10.2495/FRIAR100011, 2010.

Merz, B., Kreibich, H., Schwarze, R., and Thieken, A.: Review article "Assessment of economic flood damage", Nat. Hazards Earth Syst. Sci., 10, 1697–1724, doi:10.5194/nhess10-1697-2010, 2010.

Molinari D., Scorzini A.: On the influence of input data quality to flood damage estimation: The Performance of the INSYDE Model, Water, 9(9), 688; doi:10.3390/w9090688, 2017.

Molinari D., Scorzini A.R., Arrighi C., Carisi F., Castelli F., Domeneghetti A., Gallazzi A., Galliani M., Grelot F., Kellermann P., Kreibich H., Mohor G.S., Mosimann M., Natho S., Richert C, Schroeter K., Thieken A. H., Zischg A. P., and Ballio F.: Are flood damage models converging to reality? Lessons learnt from a blind test, Nat. Hazards Earth Syst. Sci., Discussion started: 24 February 2020, doi:10.5194/nhess-2020-40, 2020.

Scorzini, A.R. and Frank, E.: Flood damage curves: new insights from the 2010 flood in Veneto, Italy, Journal of Flood Risk Management, 10 (3), 381-392, doi:10.1111/jfr3.12163, 2017.

Thieken, A.H., Olschewski, A., Kreibich, H., Kobsch, S., and Merz, B.: Development and evaluation of FLEMOps - a new 745 Flood Loss Estimation MOdel for the private sector, in: Flood Recovery, Innovation and Response I, edited by: Proverbs, D., Brebbia, C.A., and Penning-Rowsell, E. (Eds.), WIT Press, pp. 315-324, doi:10.2495/FRIAR080301, 2008.

Wagenaar D.J., de Bruijn K. M., Bouwer L. M., and de Moel H.: Uncertainty in flood damage estimates and its potential effect on investment decisions. Nat. Hazards Earth Syst. Sci., 1-14, doi:10.5194/nhess-16-1-2016, 2016.