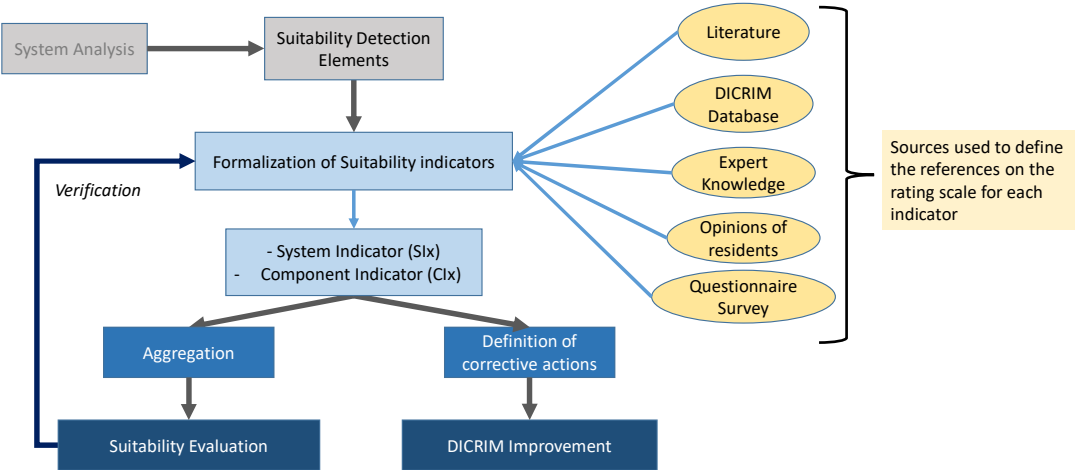


Responses to reviewer – RC2

Reviewer comment	Authors reply
<p>I encountered difficulties in reading the text, since it is not fluid, and in the understanding the workflow you used to perform your analysis. Some parts of the manuscript are, in my opinion, to be presented differently.</p>	<p>Thank you very much for your interesting comments. We hope that the changes we made following your comments will make the reading of the article more fluid</p>
<p>First of all, I did not understand how figure 3 can reflect the method used. If it is clear that a Systemic Analysis is performed for the suitability detection elements, but it is not clear to me how all the other sources used to select and identify other indicators were processed (the orange ellipses in figure 3). For example, the results of the questionnaire survey dedicated to the design of the cover page were used for the SI or for the CI indicator? I think you have used for the component analysis since the cover page is one of the document components but it is not clear when reading the manuscript. And it is the same for what you named “DICRIM database” or for the “opinions of residents”.</p>	<p>Indeed, the figure does not quite present the process that was used: first a systemic analysis to define the detection elements, then the use of 5 different sources (literature, DICRIM database...) to go from detection elements to indicators. We therefore modified the figure to show this process</p>  <pre> graph TD SA[System Analysis] --> SDE[Suitability Detection Elements] SDE --> FSI[Formalization of Suitability indicators] FSI --> SI_CI["- System Indicator (SIx) - Component Indicator (CIx)"] SI_CI --> Agg[Aggregation] SI_CI --> DCA[Definition of corrective actions] Agg --> SE[Suitability Evaluation] DCA --> DI[DICRIM Improvement] FSI -- Verification --> SA subgraph Sources L[Literature] DD[DICRIM Database] EK[Expert Knowledge] OR[Opinions of residents] QS[Questionnaire Survey] end Sources --> FSI </pre> <p>At the beginning of § 2.5 we will add a sentence: Detection elements are formalized as indicators.</p>

	The different sources are described l. 192-219. We hope that the changes made to Figure 3 will provide a better understanding of the relevance of the different sources
<p>A second problem concerns the experts: did they provide indicators or they were included in the verification phase only, or both?</p> <p>I have also encountered a problem with some terminology, it is my fault, but it is not clear to me if anchoring points on the scales, references and milestones can be considered the same concept. Please clarify better the differences.</p>	<p>Two groups of expert were involved: a first one (3 experts) worked on the formalization phase and a second one (5 experts) performed the validation tests. We will specify in the article that 2 independent groups of experts have worked</p> <p>You are absolutely right. We used several terms for the same thing and this is confusing. We will adopt a single term in the revised paper</p>
Regarding the scales of analysis, it is clear the reasons for using more than one. Conversely, I found difficulties in the understanding the way you have used the indicators for the two scale of analysis.	The assessment is carried out using the quantitative scale (0-10) but we introduced a qualitative scale (unacceptable – good) because experts look first for the qualitative class (e.g. "Poor") and then for the more precise score (e.g. "4")
Comments from Supplement document	
Please, specify what are you meaning with "fields it contains".	<p>We will give examples the fields contained in DICRIMs:</p> <p>System Indicators for the entire document and Substantive and Formal Component Indicators for the fields it contains (e.g. editorial from the mayor, DICRIM presentation, substantial information for each hazard affecting the town, emergency phone numbers)</p>
it is merely a list. it could be more interesting to add few details.	We will add some words to describe each initiative
shorter is better, deleting the first part, it sounds better: suitability of risk management materials	The title will be shortened
Please, specify what kind of model you are meaning. Document model?	<p>We refered here to the French National Model for the Application of the Environment Code.</p> <p>We will change “this model” by “this National Model”</p>

composed of: (i) an editorial from the mayor, (ii) (iii).....	We will adopt this presentation
Please add (SIx) and the number of this type of indicator	SIx and the number of SIx have been added in the new version of Figure 3
Please add (CIx) and the number of this type of indicator	CIx and the number of CIx have been added in the new version of Figure 3
It could be helpful for understanding the work flow, to add close by the arrows, the type of analysis carried on	The type of analysis carried on has be added close to the arrows (new version of Figure 3)
Is it the same of the grey System Analysis in Figure3?	Yes, it is, it corresponds to the grey System Analysis in Figure 3. We will precise this : (stemming from the systemic analysis (Ferrer et al., 2018) – Figure 3)
Please try to explain in more detail	We will give more details to explain why the various elements of description (definition, measurement scale, anchorage points...) are important to formalize indicators. We will more explain the anchorage points as references because in the grids we use the term “references” (in relation with your comment “3;5,10 - are these numbers references?”)
I have a question: on what basis did you set the scores? why does the A4 format have 5 as a score?	The scores are based on the different types of source. For instance, residents indicated that they preferred the A5 format than the A4 format, so we assigned the score 10 to the A5 format and 5 to the A4 format
3;5,10 - are these numbers references?	These numbers are references – we will more explain the role of anchorage points and references in the text above Table 1
Please try to be more clear. It is an important point in the understanding the following steps of the method. The anchoring points are the same of milestome?	We effectively used different terms to describe the same thing. We will homogenise the various terms in the article for the sake of clarity
yes, you are right, but how do you use all the five sources?	As also requested by Reviewer 1, we will introduce a table presenting which source(s) is(are) used for which detection elements
CII is the only indicator extrapolated from the survey?	CII is the only indicator extrapolated from the survey
is it a verificatioin grid?	We will change the title by Verification phase

<p>Is the formulation of the milestones subject to a subjective evaluation or did it arise from a procedure?</p>	<p>The milestones (references on the assessment scale for each indicator grid) were determined by the procedure described in § 2.5 (using several types of source)</p>
<p>In my opinion the specification of the number of years of experience is not important to say, here and in the other sentences of the manuscript. you decided he/she is an expert and the number of year is not an experience indicators, or not the only one!!!</p>	<p>As also requested by Reviewer 1, the reference to the experts' years of experience will be removed in favour of information on the reasons why we chose these experts:</p> <p>Two experts in technological or natural risk analysis were more particularly involved in questions of document content and its headings, while the third as a communication expert was involved in questions of form.</p>
<p>I'm not sure to understand the meaning. did they test the documents during an emergency?</p>	<p>“in real life conditions” will be removed. This does not add anything special and is confusing.</p>
<p>the exact name is relevance or pertinence (as in table 3)</p>	<p>We will change “Pertinence” by “Relevance” in Table 3</p>
<p>Please, explain the concept in a better way</p>	<p>The sentence refers to the verification phase and will be changed, completed and moved to the beginning of § 3.3:</p> <p>Assessing all the risk headings of a DICRIM would lead to the use of too many indicators. For each DICRIM, 112 (<i>not 114 - it was an error</i>) indicators were used because the experts were asked to evaluate two risk components (a natural risk and a technological risk defined in the measurement protocol as the presentation of the risk headings is generally homogeneous in a given document).</p>
<p>how do ou get the 114 indicators. Which indicators do they include among those described so far? 114 indicators including: 48 CI, 13 SI.....</p>	<p>We will simplify the confusing text. We have defined 48 types of indicator, some of which are broken down into several components (e.g. the colour of titles). In total, we have 112 (<i>not 114</i>) indicators: 99 CI and 13 SI. The text concerned will be altered throughout the article (§ 3.1, § 3.3 and § 4) to keep only the number of 112 indicators.</p>
<p>Please, specify the reason for which you use three verification sessions.</p>	<p>We will add a paragraph to explain why we led three verification sessions: add at the end of 1.337)</p> <p>The necessary changes were made after the first session. A further session (session 2) was conducted with the experts to check that the changes were appropriate. An improvement was achieved but some further errors in scoring or lack of clarity in the definition of</p>

	indicators were still noted. We therefore modified the indicator grids and conducted a third and final session to check the improvement of the grids (no more questions from the experts and less discrepancies in scoring between them)
Please, try to explain why the 114 indicators reduced to 90.	We will complete the sentence “Ninety indicators were assessed by the experts”: Indeed, not all indicators are evaluated for the 3 DICRIMs because some DICRIMs did not contain a component: for example, DICRIM1 did not include the municipal poster. Thus, the number of indicators evaluated on the 3 DICRIMs was 90.
where can I see your results?	We can add an appendix with 3 tables (one per DICRIM) presenting the assessment of the 112 indicators by the 5 experts
Maybe, you could first divided in factual indicators from those subjective.	The first two items in the list concern factual indicators while the third concerns subjective indicators
As an example ?? why only the dicrim 1 avarages?	The results for DICRIM 1 are presented in Table 7 as an example. We can add the results for DICRIMs 2 and 3 directly in the article or in the appendix
the weights you use come from literature or are based on your experience? please specify	These weights come from our experience
But you have results for three of them	We will add: “We have shown its feasibility by applying it to three examples”
which, however, it would be better to consult.	We can add: “which, however, could be consulted.”