Review of Manuscript NHESS – 2013 – 217

Analysis of the French insurance market exposure to floods: a stochastic model combining river overflow and surface runoff

D. Moncoulon et al.

This paper presents a methodology for the estimation of the annual insurance losses probability distribution at a large scale. It is applied here to the entire French territory. The methodology developed is based on a Monte Carlo simulation approach, and the combination of two deterministic models aiming at representing the impacts of both large rivers overflooding and local runoff. The results show that based on several strong assumptions (the objective purchased here is clearly difficult to achieve), a first evaluation of the annual losses probability distribution can be obtained for the French insurance market. This distribution shows very different features if compared to the empirical losses distribution observed in the recent period (1995-2010), which highlights the probable limited representativeness of this empirical distribution.

From my point of view, this is a clearly interesting and new contribution to the field of flood risk evaluation, that merits to be published. However, in its current version, the manuscript still includes some important weaknesses that should be corrected before publication, particularly:

- the description of the methodology frequently includes imprecise or misleading points, and also sometimes a lack of references and justification of the methodological choices.
- some important assumptions in the methodology, that may have a great impact on the results, are not really developed and discussed, especially those concerning spatial dependency.
- the conclusion is too descriptive and does not focus on the main results and learnings of the paper.

More details about these overall remarks are given below. I think an overall revision taking account for these remarks should highly improve the quality of the manuscript.

<u>About the deterministic model:</u>

- Rainfall/runoff model: this part is only descriptive with very few references, explanations and justification of the choices made in the structure of the model.
- Same remark for the river routing. Moreover, the computation of eq. 14 is not clear: many unknown outputs ($Q_t^{s1}, Q_t^{s2}, Q_t^{deb}$) for only one equation.
- The description of the damage model and its calibration is not clear: How many parameters have to be calibrated? How many events are used for this within the 1995-2010 period? Are the events presented in table 6 selected from this calibration set? Why are inf and sup values presented in this table if the model is a deterministic one? A distribution of errors seems to be built based on the calibration set (fig.7): could this distribution be validated on a validation events set? Is this error distribution integrated in the Monte Carlo simulations? Clearly this part requires to be enhanced and detailed.

<u>About the F1 model</u>:

- The nature of distributions fitted for the simulation of flood series is not detailed (this is probably done in the cited references but could be quickly reminded here). One significant

methodological limitation here seems to be the ability of distributions calibrated on short series to simulate extreme events of low probability (1/1000). This point should be discussed with maybe a reminder of the conclusions of the cited references.
- "In our approach, the river flow generated for the different stations of the same river are considered independent": Unfortunately this does not correspond to reality and seems to be a very strong assumption. The large spatial dependency in the occurrence of extreme discharge values on main rivers should clearly lead to highly increase the amounts of damages for these events. Therefore, the lack of representation of this spatial dependency should cause problems for the estimation of the tail of the damages distribution (but maybe has a lower impact on the estimation of the mean of the distribution). This point should be clearly developed and discussed.

About the F2 model:
- the spatial and temporal resolutions of rainfall fields generated are not clearly explicated: the rainfall runoff approach seems to require a high temporal and spatial resolution. But only 6h and 72h cumulated rainfall fields are evocated in the description of rainfall generation. If I well understood this 72 h time step is used for events selection, with a 2 year return period threshold on 72h cumulated rainfall. But what happens then? What is the resolution of the information provided by the rain generator and used by the rainfall-runoff model?
- "used the simulated rain fields as the efficient rain": does it mean that the ETP is not represented?
- "for each event in the event set, the analysis of the uncertainties .." Not clear. Does it mean that the damages distribution is combined with the error distribution to obtain F2?
- paragraph 4.1 indicates that only 150 years of rainfall are generated, and that "the Copula method gives us 5000 years of correlated 72 h rainfall ".. This 150 year length of simulation is not described in paragraph 3. According to this paragraph 3, the use of copulas represent spatial dependency between catchments within these 150-year generated series. How do we pass then to 5000 years?

About the combination of F1 and F2:
- The choice that is made here is to use F2, and to complement it with extreme F1 events that may better represent the tail of the damage distribution (large river overflow events). This could be more clearly explicated. Were other approaches tested? Some references on this point would be interesting.
- The description of the methodology is not clear:
    - how many years are included in F2: 1000 or 5000 (see above) ?
    - The presentation of eq 15 and 16 could be improved. If I well undertstood (but I may be wrong): Eq. 15 refers to F1 or F2 separately (thus n=1000); then F is builded by taking S equal to the maximum value of F2 (a specific variable name, for instance T, could be defined for this max. value) and selecting the $n_{A>T}$ associated events in F1; finally $n_{B>S}$ is the number of events exceeding the threshold S in F, and eq. 16 should be:

$$P_{B>S} = \frac{n_{B>S}}{n + n_{A>T}}$$

- paragraph 4.1 gives a precious summary of the overall principles of the generation of the probabilistic events set. In order to help the reader, I think this information should come earlier in the manuscript. The main part of this paragraph could be moved at the beginning of paragraph 3 for instance, before the detailed description of F1 and F2 approaches.

About the description of the Macif portfolio:

- what does the “all perils” category include? All perils, floods excepted ?
- The most important here to get an idea of the scope of this study seems to be: 1 – the national repartition of the losses between floods and all perils, and 2 – the detail of the losses in the case of floods: repartition between individual and professional risks, and repartition of the individual risks (houses, flats, etc). Other information, as the repartition of individual losses (houses, flats etc ..) in the all perils category and in the entire Macif Portfolio may be removed: this information does not appear as essential and makes more difficult the understanding of this paragraph.

Example of the Argens floodplain:

- It is a pity that the costs simulated for the june 2010 and November 2011 events are not incorporated in Table 6. This case study may illustrate the performances of the damages simulation as well as the hazard model.
- The historical claim frequency should be defined. Why a range of 1-10% should be acceptable inside the flood zone?

Conclusion:

- This conclusion sounds like a summary. Please add a paragraph providing a synthesis of the main results/learnings/perspectives of this work.