Hydrol. Earth Syst. Sci. Discuss., 9, C5960–C5965, 2012
www.hydrol-earth-syst-sci-discuss.net/9/C5960/2012/

Hydrology and Earth System Sciences Discussions

Interactive comment on "A statistical analysis of insurance damage claims related to rainfall extremes" *by* M. H. Spekkers et al.

M. H. Spekkers et al.

m.h.spekkers@tudelft.nl

Received and published: 21 December 2012

We thank the reviewer for his constructive review and suggestions. Our response:

RC1: *It seems to me though there are many other aspects of the data that could have been also studied (...). I can appreciate that the collection of the data and the presented analysis already took considerable effort, but I would recommend to state somehow in the conclusions section and/or point out in the introduction that further study could be made and that this is only a first analysis of the data focusing on a specific aspect.*

AC1: A similar question was raised in RC1 by Arnbjerg-Nielsen. We would like to add the following paragraph to the discussion section: "In an exploratory study, different damage statistics were correlated with rainfall intensity, and the number of claims was found to correlate the best with rainfall intensity. The rainfall intensity was selected as it was hypothesized to be the most critical rainfall characteristic in relationship to functioning of sewer systems. In the Netherlands, sewer systems are designed to cope with rainfall intensities up to 21.6 mm/h; therefore, theoretically, the short-duration intense rainfall events are likely to contribute to damage the most."

RC2: *As far as I could see, the analysis focuses on a single aspect: establishing a relationship between the probability of pluvial flooding occurring (i.e. the number of claims is significantly higher than can be statistically expected for claims on dry days) with precipitation at the nearest gauge station (if within a certain distance, i.e. 10 km). The statement in the abstract that a relation is established between "high numbers of damage claims ... with high rainfall intensitie" becomes clear later but is somewhat misleading at the outset, since the data are reduced to the binary outcome: no damage/some damage. A statement "probability of flood damage ... with the intensity of rainfal" would seem to me to cover better the real content of the study.*

AC2: We find this a very helpful suggestion. We would, however, rather prefer to use 'rainfall damage' instead of 'flood damage' as this is closer to that what we could possible know from the data. Thus, the second line of the abstract (page 11616, line 3) will be changed in "The aim was to investigate whether the probability of rainfall damage was associated with the intensity of rainfall."

RC3: *If the authors agree, I would suggest to point out (for instance in the conclusions and recommendations) that further study of the data could be of interest: i.e. is there a significant statistical relationship between the fraction of insured buildings* *that have been damaged and the precipitation for events that are labeled as 'damage events'; is there a significant relationship between the aggregate claim of all damaged buildings and the precipitation. These relationships would be of considerable interest for simplified flood risk calculations.*

AC3: The reviewer proposes some interesting research questions for further study. We would like to include the following on page 11628, line 9: "For simplified flood risk assessment, it could be of interest to use the insurance database to investigate relationships between the total damage of all damaged buildings and rainfall characteristics."

RC4: *The interpretation and discussion of the results appears to be focused on a 'damage forecasting' application and for such an application the results appear to be not very encouraging (as stated implicitly by the authors). As in the question above, perhaps it should be pointed out that for other applications where no real-time forecasting is needed, but only the expected frequency of floods and a measure of the flood damage needs to be estimated (i.e. to assess the frequency of total losses for a given portfolio, to examine how changes in precipation frequency due to global climat warming would affect losses) the results in Figure 3 and the very high significance that is found are quite encouraging.*

AC4: A similar point was raised in RC1 and RC2 by Kreibich. We would like to add the following paragraph to the discussion: "The results of this study is of practical relevance for insurers, water managers and meteorologists. Some insurers have indicated that the staffing of their call centres (that receive the claims) during extreme events is an issue, and that a better knowledge of what events are likely to cause considerable calls (tens of times more than on a regular day) can be helpful to adjust the capacity of their call centres. It can also be relevant for insurers when reconsidering their policy

conditions. The current 'rainfall clause' that is being used (see section 2.2) has some flaws. The rainfall intensities, for example, that are mentioned in this clause are not related to capacities of urban drainage systems. Dutch urban drainage systems are designed to cope with 21.6 mm/h; the '40 mm in 24 h' criteria, for example, should ideally not lead to any flooding. The results of this study show that short-duration intense rainfall already result in a significant number of claims. Another interesting application is the development or validation of weather alarms, which are usually based on some meteorological thresholds. Climate researchers may use the model to extrapolate probabilities of rainfall damage given some projected change in rainfall extremes."

RC5: *I would suggest to say that the logistic model in Equation (4) is only one of many possibilities, i.e. below the header of 2.5, one might say "The binary outcome ... can be linked ... using various types of logistic models. Here we choose to use ..." My doubt is not so much about the choice of the logistic function (i.e. instead of the probit function), but on the use of the precipation value I instead of its logarithm ln(I). Using ln(I) is a more logical choice, because for I=0 the condition that teta=0 is then automatically satisfied. Relating the logit(teta) to I does not achieve this and puts some unnecessary constraint on the intercept and slope parameter. Note, that if one uses the model in ln(I), the model can be also rewritten in a parametric form that is more easily interpreted: odds-ratio = (I/I_0.5)$\hat{b}$eta , with I_0.5 representing one parameter (the precipitation at which the probability of damage is 50%) and the second parameter beta determining how quickly this probability increases as I becomes larger than I_0.5. It would be interesting for instance to see whether I_0.5 and/or beta when estimated for a single rain-gauge region shows significant variations between different regions and whether these differences can be related to some region-attribute (average landuse, area, ...).*

AC5: We agree with the reviewer that different models are available for binary data.

We will rephrase the first and second sentence of section 2.5: “The outcome, damage event or not, can be linked to the maximum rainfall intensity (maximum within one day for the chosen time window z) using various types of models for binary data (McCullagh and Nelder, 1989). In this study a logistic function was used, which yields:”. We have considered to transform I by using its natural logarithm, $\ln(I)$, as it is indeed easier to interpret/visualize model outcomes. However, we would rather use the actual measured values, as the transformation may introduce an additional source of uncertainty. The model should ideally go through the origin; however, this is not entirely the case here. By dichotomizing the data according to Eq. 2 and Eq. 3, there will be, by definition, some cases that are labeled as ‘damage events’ ($Y = 1$) although it was a dry day, $I = 0$. For example, using $\alpha = 0.01$, 1% of the dry days will be labeled as ‘damage event’. It is indeed interesting to see if differences between regions can be attributed to some local characteristics of that region. See also our response to RC6 by Arnbjerg-Nielsen. However, the total number of ‘damage events’ is rather low; there will only be a very limited number of observations per region to make the analysis statistically sound.

RC6: *page 11620, line 15: “close to (not further defined)”. I had to read this a couple of time to understand. Suggestion: “at a rain gauge station that is close to (without close being precisely defined) the location”*

AC6: On second thoughts, we would rather state this sentence in a slightly different way, page 11620, line 14: “Rainfall is considered “extreme” when “rainfall intensity is higher than 40 mm in 24 h, 53 mm in 48 h or 67 mm in 72h at or near the location of the damaged property”, without “near” being precisely defined.” Mind the use of quotation marks here to, to emphasize that we quote from the report by Ministry of Transport, Public Works and Water Management (2003).

RC7: *page 11623, line 9: “maximum rainfall intensity”. “maximum” is previously defined, but for those with a short memory it might read easier “maximum rainfall intensity (maximum within 1 day for the chosen time-window Z)”.*

AC7: Good suggestion. We will change page 11623, line 9 into: “The outcome, damage event or not, can be linked to the maximum rainfall intensity (maximum within one day for the chosen time window z) using various types of models for binary data (McCullagh and Nelder, 1989).” Note that the first suggestion in RC5 is incorporated too.

RC8: *page 11625, line 9, and caption of Figure 3: “one standard deviation of uncertainty”. Suggest to add “on the empirical proportion estimate”*

AC8: Good suggestion. We will add “on the empirical proportion estimate” after “uncertainty” in line 9, page 11625 and caption of Fig. 3.

RC9: *page 11633, caption Table 2. “z= 60 mm h^{-1}”. I don’t understand this. Should it be “z= 60 min” and add “b_1 has units h/mm”?*

AC9: This is indeed a mistake. z should be in minutes, thus “z=60 min” and we will add to the caption of Table 2 that “The regression coefficient β_1 has units h/mm.”.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 9, 11615, 2012.