Review of

Characteristics of hail hazard in South Africa based on satellite detection of convection storms

By Punge et al

Summary

This article is the second revision of the paper with the same title. In this paper, the authors investigate the formation of hail events and hail hazard in South Africa as guided by 14 years of geostationary satellite observations storms. A multivariate stochastic model was built simulating event properties spanning 25 000 years of hail occurrence using overshooting cloud top detection to describe the spatio-temporal extent of potential hail events. Hail footprints were generated that could be used by the insurance industry for risk analyses.

The revised paper shows where revision was done and the authors took care to answer the questions from the first review. The suggestions made by the reviewer are minor and will help with the clarification of a few minor points.

Below are some questions and suggested corrections:

1. Page 1: "to estimate risk for the insurance sector (Punge et al., 2014; Radler et al.)." The date for the reference is missing. Please double-check check rest of the references for a similar problem.
2. Page 13-15

The description here is a bit fuzzy for me. I understand the differences in groupings of grid sizes for hail events and for time. However, the manner in which it is described is not clear for me.

For the number of events, the 0.3x0.5˚ is combined into new grids of $3x5^0$ by grouping 10x10 of the smaller grid cells (I assume the 0.3x0.5˚ grid cells) together. Then days are drawn from a newly formed $3x5^0$ grid and the surrounding 8 boxes. What are the sizes of these boxes (0.3x0.5˚ or $3x5^0$)?

Are these 8 boxes chosen in any particular manner?

The switch between the terminology of "grid", "domain" and "boxes" becomes confusing.

For time, is the centre point taken in the newly formed $3x5^0$ grid and then taking a new grid over the centre point of the size $10x6^0$

3. Section 3.1

Event lengths and widths are approximated with GEV over the exponential distribution due to the better fit. Its stated that the GEV fits well over the bottom tail of the distribution. But it does not fit well for low widths that are over-represented (over-estimated) and it does not fit well in the upper tails in that it gives unrealistic large values.

a. Which GEV function was used?
b. From the description, it is my understanding that the good fit of GEV is only for the bottom tail of the distribution for length of the storm. And not really anywhere else on the distributions? Should a different GEV function be applied to get a better fit? What statistical tests were performed to see the goodness-of-fit of the tested distributions?

c. How was the value 1.5xlargest observed value chosen as the point/place where to truncate the events? It appears this was done for the whole country – but what this checked to hold true for the whole country?
d. Figures 11 a and b and g are not discussed in the text

4. Section 3.2

The authors state in this section that due to the large uncertainty the hail size estimated were not considered for the modelling approach using geostationary satellite measurements alone. And that a severity index was created as a substitute.

The section, however, does not elaborate on how the severity index was set up in terms of the range of the scale. Or some descriptives on how this scale looks like or work in terms of the available data for South Africa. It is discussed throughout the rest of the paper but it does leave this section feeling unfinished.

Section 3.3 in terms discusses how the hail size can be calculated from the stochastic modelling using data from the ESWD and Severe Storms Archive. It makes the assumption that the largest hail size distributions over the continent will the same for South Africa. The authors can discuss the level of uncertainty (although not modelled explicitly) that this assumption can bring into the modelling process.

5. Section 3.5

- Page 21: "Even if there is a strong correlation between all regions, smaller regions tend to experience relatively higher variability"
  What is the definition of the regions in this context? And what is considered the larger vs the smaller regions?

- Page 24: The comparison between the modelled number of hail days for Gauteng (26 days) against that of Smith et al of 69 seems like a big difference. A description follows from how the results from table 3 can change when events larger than different event sizes are viewed. But it is not related back to the 69 events of Smith et al. From Table 3 it is not clear for what events sizes (>= to what cm size) the days are valid.

6. Section 4.1

Bottom page 24 discusses applying frequency-weights for Figure 17b and c but not which frequency weights are used and where it was obtained from.

- Figures 17, 18 – it is not clear if the number of events referred to are that of the observed geostationary data or from the modelled dataset

7. Section 4.2

- Page 27: “while Fig 18d presents the same occurrence for maximum hail severity indicator greater than 2.” Should this be 2cm?
- “We lso note that the local hail count per year is about 2 in KwaZulu-Natal maximum and ` in the Highveld and Gauteng region so…”
- Does this sentence refer to the number of hail events per year, hail events per year over a certain hail size/severity index? This seems like a very low value per year.
- Bottom of page 27:
- From line 465 – the event sized discussed – are these the maximum event sizes expected or the average event sizes expected per 10 year period? Where are these values compared with actually observed hail sizes as seen from newspaper/twitter reports?

Publish with minor revision that the editor can check.