A review of the manuscript
"Three variables are better than one: detection of European winter windstorms causing important damages"
by M.-S. Deroche, M. Choux, F. Codron, and P. Yiou

General remarks:

The authors aim at an identification of the most damaging events from a specific choice of 3 meteorological parameters computed from grid point values of re-analysis data over Central Europe. The approach thus aims to develop a procedure identifying potentially damaging events from the output of meteorological models (like climate change simulations), not using additional information like the exposure (distribution of vulnerable items) and neither using the spatial extension of the area affected by a windstorm nor the intensity distribution over this area. It is confirmed that the original set of 10 insurance-relevant events (used for developing the identification rules) is included in the damage prone events identified with any of the three parameters a) maximum wind speed threshold exceedance, b) relative vorticity at 850hPa and c) MSLP anomaly. The results suggest that combination of the specific three parameters used is better suited for the identification of severe insurance relevant events as there are many storms which are extreme in one or the other parameter, but only 24 storms are extreme in all 3 parameters, including the 10 most relevant ones defined from damage records.

I find it difficult to appreciate the basic approach of the paper. It seems to be physically clear that it is extreme wind speeds which are eventually causing damage, and not vorticity or pressure anomaly. Thus, there is the obvious question why the other parameters seem to add value to the simple translation of (excess) wind speed into damage, including a summation over the area hit. Can it really be assumed that exposure is irrelevant? My guess is that the requirement of additional extremes is merely a surrogate for extreme wind severity and extension which is not included in the wind speed parameter used, which limits itself to the maximum value occurring at just one grid point over the land area in the chosen Central European window. It must be explained that this is essentially the basic assumption of the approach.

It seems reasonable to assume that the maximum wind speed ratio is not a perfect proxy for the area and the integrated intensity of wind speeds. Thus, the idea of choosing just the maximum and using other parameters as secondary selection criteria is fine with me, provided it is explained why this procedure should work (Fig. 7 might give a hint), and in how far its applicability is limited to the area used. Thus, the question of the size of the area (which obviously must be discussed. The area must include both the grid point with maximum normalized windspeed over land AND the pressure minimum, plus the vorticity maximum (there could be more than one relative vorticity maximum, and even more than one pressure minimum associated with the cyclone, with only part of them located within the frame considered). Considering an area which does, for example, not include the location of pressure minima and wind maxima simultaneously will inevitably lead to failure of the approach.

The events included in the sample (of maximum wind speed only, and of the combination of parameters) are not sufficiently covered. If they are all in the list of 50 events of XWD this may would be perfect, but I expect this is not the case. Thus, what kinds of events are extreme according to the criteria, and still not produce much damage? At least show an example.

The reason for the choice of the specific parameters is also not clear. There are other parameters which could be as relevant (vertical stability and gustiness).

In summary, the manuscript needs a major revision, with the revised version clearly pointing at inherent assumptions and the limitations of the approach.

Specific remarks

Major

1. It is not clear why the Mediterranean is excluded. Shouldn't the same physics apply to Mediterranean cyclones as to North Atlantic ones? The only differences I could imagine age the influence of orography/land for the generation and the growth. These factors would, however, also be relevant for North Atlantic cyclones when they reach the continent. Presuming that the authors have also looked at the Mediterranean, I do not think it is too much

work to mention in how far Mediterranean cyclones differ with respect to the applicability of the Methodology. I think the choice of the location and size of the frame considered is of major importance for the method.

2. Some ETCs like Lothar are essentially small scale disturbances related to a larger System (Klaus, in this case). It seems to be not straightforward to find a strong low pressure core in the narrow spatial window used. Lothar considered at low resolution may be an example for this problem. Also, there is an offset between the low pressure cores and the area of maximum wind speeds. This would imply a shift of windows used for the different parameters in order to catch events properly. With the current configuration, some major relevant events may be missed. I guess this would become apparent when going beyond the 10 events listed. Is this the case? How well are relevant apart from the ensemble of 10 ETCs caught For example use the Extreme Windstorms Catalogue XWS (in total 50 events).
3. Describing the "final" methodology before providing reasoning for the chosen procedure makes this section difficult up appreciate. I thus suggest a re-ordering. Also, it should be clear throughout the chapter that the individual detection methods will be considered and compared, not just the combination.
4. Apparently, there are additional details of the methodology which are subjectively chosen to meet the 10 storms, and to exclude other extreme events. On page 4266, a gathering of timesteps into one event is described which does not just look for a maximum of consecutive time steps but involves a "simple tracking". Why? This is not in line with the idea of creating a simple procedure.
5. When combining the conditions assigned to the parameters, must they be fulfilled simultaneously? Could they as well be fulfilled at different time steps (e.g., 6, 12 or even 24 hours apart)?
6. Section 4.3 describing results shown in Fig. 7 does not point at the result that the ordering of events according to damage seems to be met by the ordering in every single parameter considered when using ERA Interim. This result is not reproduced in the reduced resolution version of ERA Interim and NCEP. You should mention the reasons for this disagreement: Are, for example, the differences between the parameters small so that small differences due to smoothing produce large differences in the ordering?
7. The agreement of the method's success between different re-analysis datasets does not automatically warrant a successful application on GCM output, as GCMs may produce a different agreement of the parameters as re-analysis.

Minor comments

1. Abstract L5ff: you are mixing different event aspects and signatures: damage, cyclone, and meteorological parameters (850 hPa vorticity is not strictly a surface signature). Please be more precise in your formulations.
2. 4258 L20: Many Atlantic cyclones do not reach Europe.
3. 4258 L22: Cyclones and windstorms are NOT the same! Such a misunderstanding, sometimes appearing under non-experts, must not be supported by a formulation which seems to accept it ("referred to").
4. 4259 L1: It is the accumulation of individual damages which causes the high loss figures for insurance, not the single local intensity and impact. This is a characteristic of the European winter storms.
5. 4260 L 11: In what sense do you want to remain flexible?
6. 4262 L 8: What do you mean with "accounts for the decrease of grid size"? Is it a reduction of resolution based on the number of resolved spherical harmonics?
7. 4263 L1: Vorticity focuses on finer structures than SLP. As the latter is computed from the former (remember the Coriolis parameter using geostrophic approximation), the information content with respect to resolution is essentially the same.
8. 4263 L8: Neu et al (BAMS 2013) also mention this issue for the cyclone tracking intercomparison project.
9. 4263 L10: Better say "remove" instead of "bring out".
10. 4263 L 19: Eq. 1 and 2 (and the respective sentence) are better at this place.
11. 4263 L 25: Leckebusch et al (MetZ, 2008) use this quantity for identifying and tracking storm fields (i.e. features). This threshold is chosen in their approach to meet the one often used for

estimating damage to buildings. Thus, it has a relevance for insurance. I would see “destroyed” as something considerably more destructive, even when it is only “partially”.

12. 4265 L 10: You apparently mean highest values, not the upper part of the plot.
13. 4268-4269: It is not clear why you speak of “likelihoods” of events being identical. Even if maxima in a parameter occur at a different time step at a different grid point in the area chosen for identificatiion, they can still belong to the same storm, as storm fields can extend over many gridpoints, even having different maxima. Thus, the requirement of “the same” time step, presumably required, is too strict. I recommend to use the time window you defined before for determining the rate of agreement!
14. 4269 L 15: I don’t understand “ordering of the distribution”.
15. Figure caption 2: Mention green points on top row, compare to points in second row.