On the clustering of winter storm loss events over Germany

Melanie K. Karremann, Joaquim G. Pinto, Philipp von Bomhard, Matthias Klawa

Replies to R. Vitolo

Dear Referee R. Vitolo,

We thank you for the comments and helpful suggestions, which contributed to improve our manuscript. Below you can find our point-to-point responses to your comments. All page and line numbers refer to the NHESSD document. The text in italics corresponds to the referee's comments; new sections in the text are marked in red.

2 Specific comments

1. page 1917, lines 23-25:
"Following the normalisation with the 98th percentile," what is being normalised to what? "the values were interpolated using distance weighted interpolation to the 0.25_ grid of the population density" Which values were interpolated? The exceedances above the 98th percentile, or the normalised exceedances (perhaps divided by the 98th percentile)? For a given population gridpoint, how many (neighbouring) values were interpolated? This delicate point needs to be clarified extremely well: I suspect that the interpolation of intense windspeeds is a tricky procedure, given the very strong spatio-temporal variability which characterises windspeeds and particularly wind gusts. Ideally, the Authors should include some discussion referring to further scientific literature on the topic.

Answer: We thank the reviewer for this comment. The text on page 1917 line 23 et seq. has been rephrased in order to clarify the method. Among other things, we highlight that we are interpolating the normalised wind values, which leads to a more flatter distributions in space than the not normalised wind. Furthermore more details concerning the interpolation method were included. Nevertheless, we did not insert a discussion about different interpolation methods as we think this would be to detailed

in this section and possible distract attention from the main storyline.
"The 98th percentile at each station is calculated for the winter half year. Following, a normalisation of the 10m-wind gust observations with the 98th percentile at each station is performed. The normalised values were interpolated to the 0.25° grid of the population density (Fig. 1c) using the inverse distance weighted interpolation of second order. This method assumes that the interpolated value for each grid box should be influenced more by nearby stations and less by more distant stations. The second order fit permits a higher weighting for nearer stations."

2. page 1919, lines 10-15: "The assignment of gridded wind data ...". It is not clear, here, what is being assigned to what. The caption of figure 1 seem to suggest that wind gusts from the DWD dataset or 10m wind speeds for the other datasets are attributed to each cell of the population density grid. This point needs to be clarified in the main text, perhaps even adding an example for a specific gridpoint, as the results might heavily (and, possibly, very heavily) depend on this choice, as the Authors themselves seem to suggest (e.g. on page 1926, lines 8-10). Indeed, losses are typically due to intense wind gusts, which have a very strong spatial variability (see the point above).

Anwer: We agree with the reviewer that the sentence was misleading. Therefore the wording on page 1919, line 10-15 was slightly changed. Nevertheless, we do not show an example in the text, as it would be to complex in this content and distract from the storyline.
"The assignment of gridded wind data (reanalyses, GCM) to each population density grid cell is done with the nearest neighbour approach."

3. Page 1920, line 5: "In this study, a new approach is used to separate events from daily maximum data:" Unclear: do the Authors mean that a new approach is adopted to identify individual, distinct loss events in the daily time series? Please rephrase.

Answer: To meet the reviewer's concern, we have slightly rephrased the sentence on page 1920 in line 5.
"In this study, the approach is adapted to identify individual, distinct events from daily

maximum data.

4. Page 1920, line 6: "The local temporal maximum of LIraw (MIraw) for a three-day gliding time window is defined as event." This is unclear. How comes that the method is able to identify storms "Vivian" and "Wiebke" as distinct individual events (compare page 1925, line 5), which are separated by 1 day, if the maximum over 3-day gliding window is used? Please explain carefully, preferably with an example from the time series (preferably "Vivian" and "Wiebke" themselves).

Answer: Both reviewers have posted comments that the description of the methods was hard to follow. In order to meet the concern of both referees and to clarify our method, we have completely revised the section about the adaption of the method to identify individual events (page 1920 line 6 et seq. – page 1921 line 1).

"In this study, the approach is adapted to identify individual, distinct events from daily maximum data. In the following overlapping three-day sliding time windows are analysed. Given that Germany is a comparatively small area, three days are reasonable for separating events. This also corresponds to the 72-hour event definition that is often used by insurance companies in reinsurance treaties (cf. Klawa and Ulbrich, 2003).

- The middle day of the three-day time window is defined as event if it is a local maximum of LI_{raw} (MI_{raw}). If no maximum is identified within the three-day window (for all $LIr_{aw} \neq 0$), the first day after a previous event (considering the last day of the three-day time window) is defined as event. For example in February 1990 Vivian and Wiebke are separated by only one day. Nevertheless, the method is able to identify both events, as the middle day of the three-day time window is a maximum (see Supplementary E).

In order to enable an accurate assignment of maximum wind values at individual grid points *ij* for all identified events the single grid points are analysed in more detail:

- For each grid point *ij*, the maximum of $\frac{v_{ij}}{v_{98_{ij}}}$ for the three-day time window is identified. If the determined maximum is not at the middle day, $\frac{v_{ij}}{v_{98_{ij}}}$ is replaced with the identified maximum value $max_{3D}\left(\frac{v_{ij}}{v_{98_{ij}}}\right)$ in LI_{raw} (MI_{raw}).

- In rare cases, events are only separated by one day (e.g. 26.02.1990 and 28.02.1990). If $max_{3D}\left(\frac{v_{ij}}{v_{98_{ij}}}\right)$ is identified between both events (here 27.02.1990), it is allocated to the event with higher exceedance of $v_{98_{ij}}$.
- To guarantee spatially coherent wind fields, larger values occurring on the first or third day only substitute the values from the middle day if multiple (spatially contiguous) nearby grid points exceed the 98th percentile."

5. Page 1920, line 15:
"For each grid point ij, . . . is identified around each event date and aggregated to the LI3D (MI3D) of the corresponding date." Unclear: what does it mean that the maximum is "aggregated" to the LI/MI? Please explain (possibly with an example).

Answer: We refer the referee to answer 4, where we explain how we modified the description of the complete methodology.

6. Page 1921, line 1:
"Only spatially coherent wind fields are accumulated to the events". This is unclear in may different ways: what does it mean that wind fields are accumulated to the events? What does it mean that only spatially coherent wind fields are accumulated, that is, how was it ensured that only the spatially coherent fields were selected? How was this all implemented concretely?

Answer: See anwer 4.

7. Sections 3.3 and 4.3: I do not understand how the calibration is carried out, how exactly the time series of daily maxima or of individual events obtained from the GCM is transformed: figure 4(d) and sentences "The correction is done by adapting the relative frequency of events per CWT in the 25 GCM simulations to the number of events per CWT in the ERAI data (see 4.3)"

on Page 1923, line 25 and
"This bias is corrected assuming the same frequency of events per CWT as in ERAI for GCM data." on Page 1926, line 25 leave unclear (to me) what is actually being done with the time series. This point is important, as one of the main claims of the

paper is that using the GCM data leads to a reduction of uncertainty, see the Abstract. Depending on how the calibration is performed, the reduction of uncertainty might even be a trivial result providing no additional information whatsoever.

Answer: Following the reviewer's suggestions, we give now in section 4.3 an example on how the CWT correction is performed.
"This bias is corrected assuming the same relative frequency of events per CWT for GCM data as in ERAI. For example, two SW events are identified for the top 30 in ERAI, which corresponds to 6.7% of all considered events. The correspondent number of events in GCM is 273 (6.7% of 4092). Thus, the top 273 SW events are included in the event set of the 4092 top events."

8. Page 1925, line 23:
"For each threshold, the selected LI samples (30, 15 and 6 events, respectively) are shown in Fig. 3." It may be helpful, especially for a reader who is unfamiliar with the return periods/ levels, to emphasise that the 5-year events are also 2-year and 1-year events: the blue bars in Figure 3 are counts of numbers which are also included in the green and red bars. In other words, "return period" should not be taken as equivalent to "intensity": the return period is a property of a set of events (not of individual events), where the set contains a mixture of events having different intensities. In this sense, I definitely agree with the spirit of the statement on page 1932, lines 9-13:

"Future work should focus on an adaption of the choice of events per return level as it could be improved by considering a mixture of events with different return levels within one winter", although I find it unclearly written (please rewrite). It might be worth to slightly expand this discussion, perhaps in Section 4.3.

Answer: We changed the wording on Page 1925, line 23 *:*
"Bar plots for different datasets and intensities (1-, 2-, 5-year return level events) are now analysed for the 30-year period."

Furthermore, we rephrased the statement on page 1932, lines 9-13 and expanded the discussion concerning CWTs*:*

"Future work should focus on a more detailed analysis of events with different return periods within one winter as this could improve results. Furthermore an investigation of the clustering within single CWTs, especially for CWTs with a high frequency of events, could be helpful for a better understanding of the physical aspects of clustering."

9. Page 1927, lines 23 and following: the Authors perform a sensitivity test by removing selected single years from the time series and refitting their models, however little comment is provided about the outcome of the test: what are the conclusions?

Answer: We included one sentence about our conclusions.
„ As the estimation of the return period is almost independent from the chosen years, the method is reliable for further application."

10. Page 1928, line 9:
"This enables more accurate estimates of the return period as well as lower uncertainties (Table 2)" How is the uncertainty in actually computed?

Answer: We thank the referee for this comment. Indeed the information in the method how uncertainty was computed were missing. We included the information in the text.
„This enables more accurate estimates of the return period as well as lower uncertainties calculated with the Gaussian error propagation (Tab. 2)."

Additionally we included the information in the caption of Table 2:
„Table 2: Estimated return periods for three different return levels (1-, 2-, 5-year) based on the Poisson distribution (Pois. RP), the empirical data for each dataset (eRP), and the negative Binomial distribution (Neg. Bin. RP; with uncertainty estimates using the Gaussian error propagation) for NCEP, ERAI and independent selected GCM samples (GCM: all runs, GCM_{corr}, 37 ESSENCE runs: ESS_{corr}, 3 20C runs from MPI: $20C_{corr}$, PRE_{corr} from MPI, 3 CSMT runs from MPI: $CSMT_{corr}$; all runs indexed with corr are bias corrected based on CWTs) considering only the number of years available for each dataset respectively. The number of years is indicated below each dataset. For further details see Table B1.

3 Minor points

1. Page 1925, 1st line of Section 4.2: Figure 3 is not a histogram, it is a bar plot.

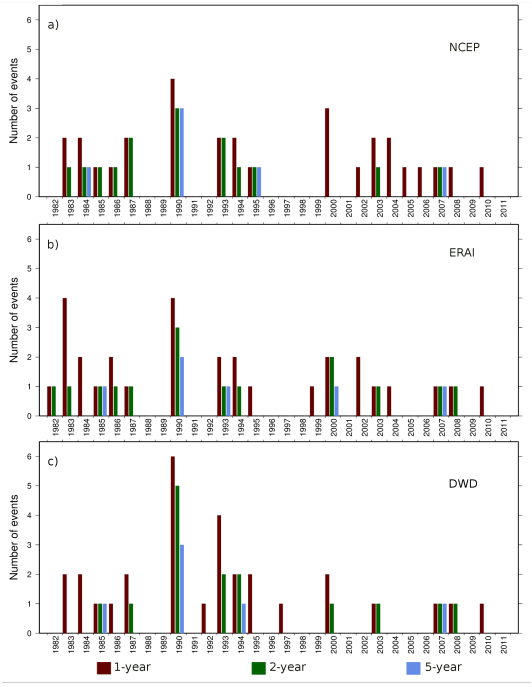
Answer: We changed the wording from histogram to bar plot.

„Bar plots for different datasets and intensities (1-, 2-, 5-year return level events) are now analysed for the 30-year period.“

2. Figure 3: it might be beneficial to use one and the same range [0;6] on the y-axes, to highlight the similarities and differences between the three datasets.

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., 2, 1913, 2014.

Answer: As suggested by the referee, we changed the range of Figure 3a) and b) to the same range as 3c). The new figure is as follows:



Supplementary E:

Date	LI_{raw}
20.02.90	0
21.02.90	0
22.02.90	0
23.02.90	0
24.02.90	0
25.02.90	0
26.02.90	**270**
27.02.90	31
28.02.90	**281**
01.03.90	158
02.03.90	0
03.03.90	0
04.03.90	0
05.03.90	0

Time series of LI_{raw} between 20 February 1990 and 05 March 1990 based on NCEP data. Identified events are marked in bold.