

Interactive comment on “Uncertainties in Forecasts of Winter Storm Losses” by Tobias Pardowitz et al.

Tobias Pardowitz et al.

tobias.pardowitz@met.fu-berlin.de

Received and published: 6 October 2016

Reply to RC1

The authors would like to thank the reviewer for the very thoughtful review. The comments, suggestions and numerous technical corrections will certainly help us improve the manuscript. Point-by-point answers (in *italics*) to the comments are given in the following.

Individual scientific question/issues (“specific comments”)

Title: Could better reflect the study for example 'An Analysis of Uncertainties in Forecasts of Winter Storm Losses'.

New title based on both reviewers' comments: “An Analysis of Uncertainties and Skill

C1

in Forecasts of Winter Storm Losses”

Page 1 line 12 (abstract): 'Deterministic assessment of damages' please clarify, is this prior to the event? Even knowing local vulnerabilities would not allow for a yes/no damage assessment until after the event when the damage is done.

We exchanged “..deterministic assessment of damages..” by “..deterministic prediction of damages, even if the forecasted wind speed contains no uncertainty”

Page 3 lines 8-11: It is stated that 'hail induced damaged cannot be separated...this poses another uncertainty that needs to be reflected in the relationship between local winds and resulting damages' however this is not mention again in the paper. How this could be done and what it may show?

This is a very interesting and tricky point! From a data point of view, hail induced damages cannot be distinguished from wind related damages in the dataset we use. According to the provider of the dataset (GDV), winter months are dominated by wind-storm damages while summer is dominated by hail induced damages. However in rare cases of severe winter storm events, hail damages may occur. E.g. it is known that damaging hail occurred during the frontal passage of storm Kyrill in 2007 (Fink et al, 2009). Taking into account the occurrence of hail and resulting damages could be done based on additional predictors such as the “convective available potential energy” (CAPE) and “convective inhibition” (CIN). Based on a logistic regression model with multiple predictors, both the individual effect of hail but also the contribution of hail in the case of winter storms could be quantified. It can be assumed, that the probability of hail will increase in case of the most severe winter storm events. Thus for high wind speeds, the damage probability forecasts (which neglect the effects of hail) might be underestimated. Considering the reliability diagrams for the probabilistic forecasts (not shown) we do find such underestimation of the probability forecasts. However a more in depth analysis is needed to clearly attribute this to effects due to hail. This has not been the scope of this paper but we plan to address this in further research.

C2

Furthermore we will add the discussion outlined here to the discussion section within the manuscript.

Page 5 line 12: How can post-processing help adjust the spread of an ensemble? Can this be made more explicit/clearer?

We added clarification by modifying the original quote (Page 5, lines 10-12):

“Despite of such sophisticated techniques for the perturbations, ensemble forecasts still often tend to an under-dispersion. A post-processing can help to adjust the spread of the ensemble, what is meant in this study by the term ‘calibration’.”

To:

“Despite of such sophisticated techniques for the perturbations, ensemble forecasts still often tend to be under-dispersive. This means, that the spread of the ensemble members (the members being discrete random draws of the forecast PDF) may be too small and it may not reflect the full uncertainty inherent to the forecast. ‘Calibrating’ the ensemble spread, which is part of sophisticated post-processing techniques can thus help address such under-dispersion of ensemble forecasts (see Broecker et al.,2008).”

Page 5 line 20: Can the size of the Gaussian kernel be specified?

The ensemble post-processing is performed on a grid-cell basis and the Gaussian kernel determined individually for each grid-cell. Additionally they are dependent on the actual forecast. The kernel size can thus strongly vary depending on grid point and specific forecast situation. Further details can be found in Broecker, J. Smith, L, 2008. For clarification, we will comment this in the manuscript.

Page 8 line 10: Clarify what 'low forecasted gusts' are.

To clarify, we revised the sentence to:

“.. is stronger for weak gusts (<5m/s)”

C3

instead of:

“..occurs mainly for low forecasted gusts”.

Page 8 line 20: Clarify this sentence, how is a doubling shown? The sentence and Figure don't appear to correlate.

The original quote read:

“Considering the Talagrand diagrams for the post processed forecast (Fig. 2, right) shows nearly equally populated ranks. Slightly higher populations (doubled in case of forecast lead time of 1 day) are found for the lowest and highest ranks.”

Figure 2 (right) shows a rather flat Talagrand diagram. For most ranks a frequency of slightly less than 0.025 are found. Considering forecasts of lead time 1 (red), the ranks 0 and 50 are populated with a frequency of about 0.05 which is about doubled compared to the other ranks. For longer lead time the overpopulation (underdispersion of the ensemble) is smaller which is why we generalized the statement that for the lowest and highest ranks slightly higher populations are found. We thus think that this statement does correlate to the presented Figure. However, to clarify we revised the above written quote to:

“Considering the Talagrand diagrams for the post processed forecast (Fig. 2, right) shows nearly equally populated ranks. Slightly higher populations are found for the lowest and highest ranks. In case of forecast lead time of one day (red), the lowest and highest rank are populated with a frequency of about 0.05 which is roughly twice the frequency found for the intermediate ranks.”

Results section: Over-forecasting is not mentioned. On page 9 lines 8 and 9 '10-20% probability using full uncertainty in most of the affected areas' is stated. I understand the nature of probabilities however it would be good to have a comparison between forecast probabilities for the loss ratios and in the areas where damage was observed i.e. how often is 10-20% forecast and damaged observed?

C4

This is a very good suggestion. We will add a discussion on the reliability of the probabilistic forecasts. For this, we will add a figure [see additional Fig. 1] to the manuscript containing exemplarily the reliability diagrams for the forecasts with lead time 3 days. Compare Wilks (2008), Chapter 8 for details on reliability diagrams.

In case of the deterministic forecasts (black circles) they show, that in about 3% of all cases for which the forecasts reads “no event” a loss event has actually been observed. Similarly in about 97% of the cases for which an event is forecasted a loss event actually occurred.

Considering the probabilistic forecasts we find, that if forecasted probabilities are low (<5%), very few events are observed. Using the forecasts treating both uncertainties and using the dressed ensemble results in an observed relative frequency of only 0.4% (which is a considerable improvement compared to 3% in case of the deterministic forecast “no event”).

The diagrams show that in general an under-forecasting occurs for the probabilistic forecasts. Over a broad range of probabilities the observed relative frequency of events is found to be considerably higher than forecasted. As an example, one might consider all cases in which the forecast probability reads 30%. According to the reliability diagram, in 40-60% of these cases (depending on which uncertainties are treated) an event had been observed.

The diagrams furthermore show that this under-forecasting is successively reduced (and thus the reliability increased) by explicitly treating the different uncertainty sources. Particularly for intermediate forecast probabilities, the distance of the reliability diagram to the diagonal (representing the optimum: forecast probability = observed relative frequency) is reduced and is lowest if both uncertainty sources are treated. Also, it can clearly be found, that the reliability is further increased when using the dressed ensemble instead of the raw ensemble forecasts.

Page 9 line 10: What is meant by 'objective verification'?

C5

To be more specific, we revised the statement

“by means of objective verification”

to

“by means of Brier Score and Brier Skill Score, which are objective measures for the quality of probabilistic forecasts.”

Page 10 lines 8-14: Could anything else be said about Fig 5? Perhaps correlate BSS to orography? Comment why the BSS in some areas is so low at -10? Could Fig 5 be related to the forecasts in Fig 3?

Our results show distinct variations over Germany. Even though we did not calculate this, we know that there is a positive correlation. However, technically there are several further questions when correlating the BSS results (on district level) for example with a gridded orography which itself varies across the district. Thus we do not think there is further gain of information by correlating BSS and orography. The BSS is calculated as $1 - BS/BS_{ref}$ (compare eqn. 2). Consequently, the BSS potentially ranges from 1 (perfect forecast) to minus Infinity (perfect reference). Since BS_{ref} in our study is rather small in most cases, negative BSS tends to quickly grow very large. The relevant information in this case is actually that the BSS is negative (i.e. a worse than the climatology). The scale is thus chosen from -10 to 0 to ensure that all negative BSS values are colored in blue. This does not mean that BSS values are -1 though. We will simply remove the -10 from the scale to avoid this confusion!

Page 10 line 5: Don't agree with the last part of this last sentence, it depends on the meteorological situation.

The reviewer is right, the underdispersion is different depending on the meteorological situation. On average we find a stronger underdispersion at short lead times, however we will modify our statement reflecting the reviewers point.

Page 10 line 15: Clarify 'this effect' (this refers to Page 11 line 15)

C6

We agree that the sentence needs some clarification. We rewrite the sentence “As stated, for short lead times and low impact situations this effect due to an explicit and full uncertainty treatment is negligible.”

to

“As stated, for short lead times and low impact situations the effect from a full uncertainty treatment is negligible.”

Page 10 line 25: Could ‘how these findings could potentially be used in risk based warning systems’ be expanded? More explicitly how could forecasters use this approach to improve forecasts and warnings?

The original quote (Page 10, lines 13-25):

“Overall this study shows, that skilful predictions of storm loss occurrences on lead times of several days can be made using the presented (fully probabilistic) framework to integrate meteorological forecast uncertainties and uncertainties resulting from impact model. Findings thus demonstrate their potential use in risk based warning systems.”

is expanded to address the reviewers point(s):

“Overall this study shows, that skillful predictions of storm loss occurrences on lead times of several days can be made using the presented (fully probabilistic) framework to integrate meteorological forecast uncertainties and uncertainties resulting from impact model. Such quantification of both, potential impacts of severe weather and their respective likelihood forms the basis for developing risk based warning systems. By quantifying impacts and their likelihood, which is particularly relevant to recipients, the acceptance of weather warnings might be strongly enhanced. As one of the first national weather services, the UK Met Office has recently moved on to a risk based warning system (Neal et al., 2013). The basis of such warning system is formed by the risk matrix, composed of the two dimensions impact and likelihood. By quantification of both these dimensions, the presented framework can thus directly feed into such a

C7

warning system.”

Reference: Neal, R. A.; Boyle, P.; Grahame, N.; Mylne, K. Sharpe, M. Ensemble based first guess support towards a risk-based severe weather warning service Meteorological Applications, John Wiley Sons, Ltd, 2013

Technical corrections

Generally: Define when talking about mean wind speeds and wind gust speeds. For wind impacts and studies using wind data knowing the difference is important.

Ensure paragraphs aren't one sentence long.

Keep consistent phrasing e.g. 'damage uncertainty' not 'impact model uncertainty' (e.g. page 2 line 14); 6-hourly (page 4 line 1) and 6 hourly (page 3 line 16).

Be consistent in the use of 'damage modelling uncertainty', 'meteorological uncertainty' and 'full uncertainty/both uncertainties' e.g. Page 9 line 1 it's referred to as 'damage occurrence uncertainty' and page 9 line 5 vs Figure 3 and Figure 4.

Ensure consistency of Brier Score and Brier Skill Score capitalisation or define then use BS and BSS don't mix.

Consistency needed for 'winter storm Britta' references including removing capitalisations.

Change all references of 1d, 3d and 9d to 1 day, 3 days and 9 days respectively including in figures.

Keep 'low' and 'high' threshold wording consistent, 'lower' and 'higher' thresholds are referred to on page 9.

Check references in particular comma placements and journal titles.

Page 1 line 14: clarify what are the 'two individual contributions'

Page 1 line 15: 'forecast skill' of what?

C8

Page 1 line 17: define 'district level'
Page 2 line 8: source 'of' uncertainty not 'for'
Page 2 line 17: approach 'when' predicting
Page 2 line 26: comprise 'of' daily data
Page 3 line 2: 'insurers' not 'insurances'?
Page 3 line 15: define DWD
Page 3 line 16: 10 meter's' above 'the' ground
Page 3 line 18/19: 'hourly 10m wind gusts' instead of 'hourly wind gusts in 10m height'
Page 3 line 20: 'maximum 10m wind gusts' instead of 'maxima of wind gusts in 10m height'
Page 3 line 22: ECMWF 'has' operationally 'run' its
Page 3 line 24: can remove 'the'
Page 3 line 25: check sentence structure and TL use
Page 3 line 28: 'In' December or re-arrange sentence
Page 4 line 1: remove space between 10 and m
Page 4 line 11: 436,905 not 436.905; comma needed after 'equation'
Page 4 line 15: TL not TL
Page 4 line 18: maximum 10m 'wind' gusts
Page 5 lines 6 and 7: Makes very little sense, please revise
Page 5 line 11: 'Post-processing' no 'A' needed.
Page 5 line 12: 'what is meant in the study by the term 'calibration'? please revise

C9

Page 5 line 13: depending 'on' the, not 'of'
Page 5 line 25: What are CDF's?
Page 5 line 26: PDF already defined
Page 6 line 6: move comma to after district and remove from after derived
Page 6 line 18: no 'the' needed before Sec. 3.3
Page 8 line 8: 'range is only between'
Page 8 line 15: 'previously' not 'above'
Page 8 line 20: Don't need 'Considering' at the beginning of the sentence.
Page 8 line 28: add 'observed' to make 'the observed loss ratio'
Page 8 line 30: add 'only' after 'uncertainty' and clarify/reword the sentence.
Page 9 line 1: 'uncertainty only for winter storm Britta' or 'uncertainty in the case of winter storm Britta'
Page 9 lines 2-4: reorder to refer to both uncertainties then dressed ensembles, this fits better with the order of Figure 3.
Page 9 line 9: 'forecasts' not 'forecast'
Page 9 line 13: 'When averaged over' not 'In average over'
Page 9 line 19: '3.5), with the climatology as a reference forecast, it is confirmed that'
Page 9 line 23: 'For a lead time'
Page 9 line 30: define 'high impact events'
Page 9 line 30: 'Even for a lead time'
Page 10 lines 4 and 5 and Page 11 lines 8 and 17: not sure 'stronger' and 'strong' are

C10

the correct words, please revise

Page 10 line 6: 'in the case'

Page 10 line 16: 'winter storms' instead of 'severe winds'

Page 10 line 24: 'In forecasting'

Page 10 line 30: 'too little ensemble'

Page 10 line 2: 'for short lead times'

Page 10 line 3: no comma needed after 'fact'; 'uncertainties are smaller at' no need for 'obviously'.

Page 10 line 7: 'situations'

Page 10 line 8: 'improvement in forecast'

Page 10 line 16: 'large lead times'

Page 10 line 23: 'Overall, this study shows, for this particular case study, that'

Figure 3: Units on right hand side, referred to in

Figure 4: Make scales the same so easier to compare. Caption could say 'Lead time dependent Brier Skill Score (BSS; employing climatology as the reference forecast) for events with loss ratio exceeding low threshold (0.0001o/oo) (left) and loss events with loss ratio exceeding high threshold (0.001o/oo) (right). Shown in black symbols are verification results for the four different set-ups, red triangles show verification results using the ensemble dressing post processing system.'

Figure 5: 'for loss ratio exceeding low threshold (0.0001o/oo)'

Thank you for these technical corrections! We will take care to incorporate them into the manuscript.

C11

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., doi:10.5194/nhess-2016-182, 2016.

C12

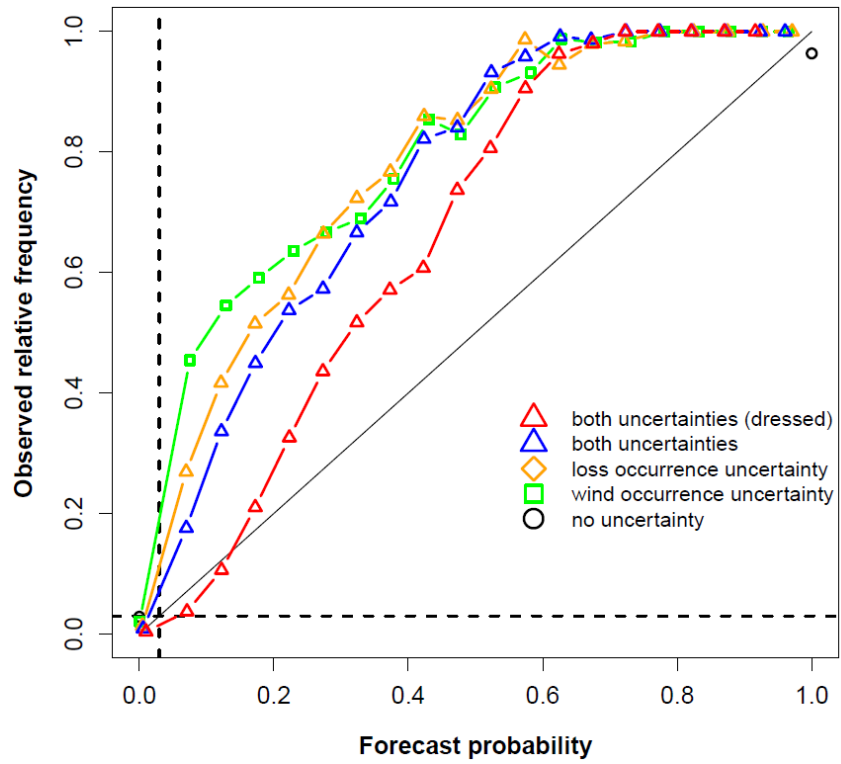


Fig. 1. Reliability diagrams for the forecasts with lead time 3 days for the high loss threshold (0.001‰).