The article describes DWD's system to postprocess ensemble forecasts using the classical techniques linear and logistic regression with a focus on severe weather. While there is nothing new in the statistical methodology, the most interesting part lies in how the methods are applied to the data. The descriptions of the selection and design of predictors, how predictions are made at any location, how to deal with model changes, rare events etc. are all of interest to readers involved in operational forecasting. Possible weaknesses of the system could be discussed in more detail along with prospects of using more modern statistical methods. Please see below for more specific comments.

Specific comments

- line 63: Not sure I understand what the meaning of "... in order to avoid underestimation of forecast errors on longer time scales"
- line 106: As I understand, the combination is done after calibrating each model, but there seems to be no information about how this is done, apart from the citation of Reichert et al. (2015). Maybe a short paragraph on this interesting issue could be added to section 2.
- line 110. I find "Ensemble-MOS" a bit confusing here. Maybe it could be added that it is the name of DWD's system rather than referring to all "EMOS" methods.
- line 135, Fig. 4d: Why do you use absolute errors instead of squared errors which is both more precise and commonly used?
- Fig. 4d: What is meant by "forecast of absolute error (m/s)"? Do you make predictions of the absolute errors directly or do you compute the absolute error after the gust forecasts are made?
- line 140: Is the upscaled product an ensemble of 10x10x20 members or 20 members?
- line 145: What kind of bias metric is this, given that the forecasts are probabilities exceeding 0.1 mm/hour?
- lines 160-165: Here results using Ensemble-MOS are presented before Ensemble-MOS is described in section 3. This is unfortunate and can easily cause misunderstandings, e.g. it is not obvious that Ensemble-MOS produce deterministic forecasts. Please consider moving this paragraph.
- line 167: As I understand, it is also used to make deterministic forecasts for continuous variables. Since this is a bit unusual, it should be stated explicitly somewhere, the sooner the better.
- lines 172-174: An alternative explanation: linear regression applied individually to an ensemble results in an ensemble of conditional means which is not the same as the conditional distribution of the predictand/observation given the predictors. In practice, the more uncertainty there is, the more will this adjusted ensemble underestimate the uncertainty.
- line 193: Is it predictands or predictors? The sentence is unclear.
- line 195: Would "binary predictands" be more precise than "probability predictands"?
- lines 195-200: Are then predictions from the linear regression interpreted as probabilities of exceeding the given threshold? Is there any statistical justification of

the approach? Any references? How good is this approach compared to logistic regression? Maybe this paragraph should be expanded somewhat.
- line 202: replace predictor by predictand?
- line 237: I do not understand the sentence "Modeling this error ...". Do you mean that separate linear regression with the absolute error as predictand are made to predict the absolute error?
- line 240: Any reference or better?
- line 243: Do you mean the regression coefficients are interpolated separately?
- line 254: I would claim that logistic regression is no longer state of the art, at least in the research literature.
- Equation 3: Maybe it could be added that conditional independence is assumed
- lines 276-277: Not sure I understand this sentence. E.g. if there are 30 events in the training set, you would restrict the number of predictors to three?
- line 299: These are deterministic predictions of gusts, or?
- lines 300-301: Have you considered adding predictors describing the sites, e.g. elevation, exposure, land type etc.?
- lines 307: please consider using "deterministic wind gust forecasts"
- line 311: have you investigated whether the globally fitted logistic regression make well-calibrated probabilities for each station, say, for the lower thresholds to get sufficient data?
- lines 312-326: these paragraphs need more accurate explanations. What is the relation between logistic regression and the logistic distribution? What is the fitted logistic distribution used for? What is the purpose?
- line 400-402: what about raw ensemble probabilities for logistic regression?

Technical corrections

- line 40: Is "landmarks" appropriate?
- line 57: Should it be postprocessing?
- line 60: "coupula" should be "copula"
- line 62-63: maybe just "... tailored to postprocessing ensembles for extreme and rare events"?
- Figure 2: What is "MA EMOS"? Not easy to understand "Ensemble-MOS forecast of absolute error of Ensemble-MOS" either.
- line 202: each predictand(?), ...
- Equation 4: Parentheses are missing(?). Summation should be over both terms, I presume.
- line 299: stepwise linear regression(?)
- figure 4: text in the plots is too small. What are the dashed lines?
- figure 6, caption: "is between"