Comparison of two physical- and statistically-based post-processing methods for high-resolution NWP visibility forecasts over 13 North European airports

Nikkilä, M., Ylhäisi, J.S., Neiglick, S., Kolvunen, T.
Finnish Meteorological Institute jussi.ylhaisi@fmi.fi

Short-term visibility forecasting for Terminal Aerodrome Forecasts (TAFs) is one of the remaining challenges where human forecasters are still able to use their expertise to outperform direct model output (DMO) or many of the solely model-based post-processing products. In this study, we apply two distinct post-processing methods for 24 hour TAF visibility forecasts over the Northern European domain covering 13 airports and compare their forecasts with the Direct Model Output (DMO) and the Aerodrome METAR model (AROME).

• The first post-processing method is a physical-based diagnostic method, which has been developed and empirically refined by the duty forecasters at the Finnish Meteorological Institute over the course of several years. The method takes into account the worsening of the visibility in the default case, caused either by precipitation or by an increase in the distance between the sun and horizon. In addition, the analogue forecasts do not show a similar degradation as a function of forecast length and the additional benefit of having access to observational information is not any more evident from the forecasts: Even the 3-hour forecasts perform very poorly as compared with DMO. Analogue method could be externally refined through using different weighting schemes, but unfortunately this kind of optimization was computationally too demanding due to long time series used in the calculation of analogies. Physical based post-processing method clearly shows more promise to it, which has also been observed in operational setting. There remains plenty of room for the forecaster to make subjective evaluation of the available information.

Figure 1. Airports analysed in the study

Methodology

Physical-based method

The development of this physical-based diagnostic method is mainly inspired by the characteristics of the 8515 model (Borgström et al., 2015) and the accumulated experience of the aviation forecasters in using it. HARMONE-AROME has been implemented and widely used for effective aviation weather forecasting at FMI, even though its success in HARMONE-AROME has largely replaced it, however, the latter model simulated visibility values were found to be very consistent with the actual observed values. The model uses a fuzzy logic. In addition to allowing a degree of similarity to each observation, the model applies weights varying for data points, depending on the METAR variables and the difference of the similarity compared to present time (the most recent observation is given the highest weight). Finally, the most similar weather situation from the past is used for forecasting. The used method is comprehensively described in Tuba and Börjesson, 2017.

Analog method

The analog method does not need any model forecasts as it is a purely observational, time-series based forecast method. In the analog method, past conditions that are similar to those of the forecast are identified and weights of data points are then calculated to identify the closest similarity of the present time with the closest past time. The present time is then weighted with weighting factors, depending on the METAR variables and the difference of the similarity compared to present time (the most recent observation is given the highest weight). Finally, the analog method is a simple and robust approach which is able to be used for the entire domain. The presented method is comprehensively described in Tuba and Börjesson, 2017.

Results

The verification results in Tables 3-4 are calculated for the Sep/Oct 2017 period, based on observational verification scores and using NOAA Aire 3 visibility classes. The verification scores are Probability of detection (POD), False alarm rate (FAR) and Pearson’s skill score (PSS). Since we aim to see how the both the post-processing methods are improving over the DMO output, but the physical method performs much more consistently over the whole visibility distribution, the physical method is not really able to produce values >.150. Even though these events constitute a minor fraction from the observations, they are still a significant fraction compared to DMO for all classes and scores. If the definition used in DMO values are extended to include all of the independent two classes, improvement of physical method is not as evident as the presented results show. However, the analogues in terms of the relative values show almost no changes from the past two model versions. These results show how the both the post-processing methods are improving over the DMO output, but the physical method performs much more consistently over the whole visibility distribution, the physical method is not really able to produce values >.150. Even though these events constitute a minor fraction from the observations, they are still a significant fraction compared to DMO for all classes and scores. If the definition used in DMO values are extended to include all of the independent two classes, improvement of physical method is not as evident as the presented results show.

One of the main motivations for the development of physical visibility scheme has been the observation that models cannot properly simulate local phenomena which are observed in the visibility observations. In cases where the weather situation is seen to be more related to the production of physical mistakes. An open-source C=Cloud repository is available for the visibility parameterisation and >70 other diagnostic model variables is available at https://github.com/finlabs/visibility/models/parameterization

References
