Manuscript Review: W-band Radar Observations for Fog Forecast Improvement: an Analysis of Model and Forward Operator Errors

Thanks to all of the reviewers who kindly gave their time to analyse the article and returned some relevant comments to help clarify the research for future readers. We hope to respond to your queries below.

Reviewer 2

Comment: In the end, the authors have NOT yet assimilated radar reflectivity and demonstrated the improvement in fog forecast. Therefore, wording like "after selecting the best background profile, a good agreement was found between observations and simulations" in the abstract is really misleading. What the authors show is the agreement between observations and "selected background profiles", which is not surprising because the "best background profile" was selected using observations as a reference. If I somehow misunderstood the manuscript and if new simulations were indeed performed using the best background profiles, then this leads to an even bigger issue. By definition, the prior is NOT supposed to see the observations beforehand. Therefore, if new simulations were performed, they must be performed for a different case or time period, and I don't see any other cases different from those listed in Table 3. This kind of misleading statements can be found in Section 4.4 and Conclusions as well, which needs to be more precise.

Response: This is a good point- we have not yet shown that the background profiles are the best profiles for retrievals and therefore the wording in sentences such as these is too strong. To clarify, forecasts from the AROME model were not run again. The profiles were simply selected from the values of their simulated reflectivity. We expect that a better selection of background profiles will improve the data assimilation of radar reflectivity and therefore the fog forecast in future studies following this first part.

As you said, it is important in classical data assimilation techniques that the prior does not see the observations. However, for cases (such as we have shown for fog forecasts) where temporal and spatial errors can be significant, these techniques can also be ineffective (Ravela et al., 2007) and thus require an extra step before classical techniques can be applied. The method that we have shown in this paper therefore made use of a method to select a background profile which could account for temporal and spatial errors, without directly changing the background profile to suit the observations. In that sense, the selected background profile is still a forecast independent from the observation.

Change: Line 18 re-worded to 'After selecting the background profiles with the best agreement with the observations, the standard deviation of innovations (observations - simulations) was found to decrease significantly'

Line 432: 'Using the MRP selection, simulated reflectivity showed better agreement to observed reflectivities with the choice of a more appropriate background profile.'

Line 498: 'This study shows that, after removal of the largest background errors, the forward operator used in this study is able to replicate similar values of radar reflectivity from the background profiles, compared to the profiles observed during fog conditions'

Line 552: When a better agreement was found between the background profile and observation, the radar simulator was also found to be suitable to simulate the BASTA cloud radar reflectivity during fog conditions paving the way for larger model evaluations during fog events

Comment: Additionally, I think the use of "Innovations" is too strong and not accurate. It is an improvement, not an innovation.

Response: Here the term 'innovation' refers to observation minus simulation from the background profile ($y - H(x_b)$) values in the field of data assimilation.

Change: I've clarified the definition when the term is first used.

Comment: "Visibility measurements were averaged over 10-min period", meaning for both observations or simulations or both?

Response: This was only done for the observations, to account for noise – the model outputs were only available at a resolution of once per 10 minutes.

Change: Added: 'As model outputs were available with a temporal resolution of10 minutes, these were not averaged'

Comment: Descriptions in Section 3.2 are quite confusing to me, and I am not sure that readers are able to replicate results. What is "fog profile"? "Visibility measurements were averaged over 10-min period", meaning for both observations or simulations or both?

Response: It was used to describe a 10-minute time block where the model or observations were under fog conditions. I agree that this wording is not totally clear, so it has been re-written to clarify to the readers.

Change:
A comparison of observed fog to fog predicted in the model- for the time and grid point corresponding to the time and location of the observation- was carried out. Visibility measurements, taken from the DF-320 visibility sensor, were averaged over a 10 minute period, and where visibility values of lower than 1 km where observed, this was considered as a fog 'block'. The same threshold was used with visibility diagnosed from the model to define model fog 'blocks'. As model outputs were available with a temporal resolution of 10 minutes, these were not averaged. The accuracy of the model was then analysed by comparing each 10-minute block in the model against each block from the averaged visibility. Observations where rain was sensed with the rain gauge and simulations in which rain was present in the bottom layer were not considered as fog. The commonly used contingency table based on this comparison is shown in table 3 where GD indicates cases of good fog detection, FA cases of false alarm, ND cases of missed fog events by the model and CN correct negatives.

Comment: How might a choice of 28 km x 28 km domain relate to the sample size of 15248 in total in Table 4?

Response: As a 'fog block' in the model was diagnosed just from the grid point corresponding to the SIRTA observation site, the domain does not relate to the sample size mentioned in table 4. Instead, the sample size corresponds to one grid point every ten minutes for each day between 2nd

November until the 19th of February, with some blocks missing due to missing data, in which case no comparison was made.

Comment: The manuscript will read better if things are defined and clearly stated in a slightly different order. For example, how to define fog thickness in observations AND simulations? The term is introduced in 3.1 (page 7) but is not defined/explained until page 11. Even so, it is still unclear how exactly it is done and if it is the same for both observations and simulations… It would be nice to mention that earlier, so readers can connect Fig. 4 and the all exercises/results better.

Response: Thanks for the comment. We have tried to clarify what is meant by fog thickness in section 3.1. We did not want to discuss in depth how fog thickness is derived from the model in section 3.3, as in this section we do not discuss the comparison with observed fog thickness. However, when comparing model fog thickness/ fog top heights to observed fog thickness/fog top heights, in section 3.4, the way the fog top height is defined is important for the reader, and so we think it is more relevant to keep the explanation of fog top height prediction by the model in this section.

Change: Added in 3.3 (page 9) **‘**The fog thickness was diagnosed from simulated reflectivity values and is explained in more detail in section Section 3.4’

Added in line 290: Fog thicknesses were derived from the radar observations during fog conditions. This was found from the height at which the radar reflectivity dropped below the larger of −45 dBZ or the sensitivity of the radar (whichever value was greater) at that range gate.’

Comment: Another example is the information on parameter ranges on Page 15.

Response: In response to this comment and larger concerns of the other reviewers, the discussion of the parameters on page 15 has been significantly reviewed.

Change: Section 4.1 rewritten. Lines 325-369:

In the pair of equations, N(D) is the droplet number concentration where D is the droplet diameter. Coefficients a and b determine the mass-diameter relationship of the droplets, which, when applied to cloud droplets are well known due to their spherical nature, and are set at 524 and 3 respectively. α and ν are fixed coefficients referred to as the shape parameters and are set to 1 and 3 respectively in ICE-3 for cloud liquid droplet over land. N_0 is the total droplet concentration and is set to 300 in ICE-3 for liquid cloud over land. M is the liquid water content of the grid point in $kg.m^{-3}$.

The advantages of using this modified gamma distribution are that the shape and median diameter of the distribution are modified with the liquid water content and number concentration of the cloud. For example, when using the modified gamma distribution with a total concentration of $30cm^{-3}$, the median diameter will be greater than for a total concentration of $300cm^{-3}$, as illustrated in figure 5.

As all parameters of the modified gamma distribution except for the liquid water content are held constant in ICE-3, when radar simulations are made for cloud with a droplet size distribution which the parameters do not accurately describe, errors are likely to be made in the calculation of radar reflectivity. In order to assess this uncertainty, simulations were made on an AROME model profile in fog conditions, for which the size distribution parameters were perturbed. These perturbations would need to reflect potential variabilities seen in (continental liquid water) fog and low liquid cloud.

Microphysical observations have been investigated on fog events in previous works(Mazoyer et al., 2019; Podzimek et al, 1997) which tend to show lower droplet concentrations than is prescribed for continental clouds in the ICE3 microphysical scheme (of 300cm^{-3})). From the works of Mazoyer (2016), which looked at median droplet concentrations for continental fog events, and Zhao (2019), which investigated the microphysics of continental boundary layer clouds, reasonable lower and upper bounds of the N_0 parameter of 30 and 300cm^{-3} were decided. Figure 5 shows the difference in cloud droplet distribution shapes when these two values are used.

As the α and ν parameters both affect the width of the size distribution (as may be seen in figure 5), it has been a common approach (Mazoyer, 2016; Geoffroy et al. 2010) to fix α and to optimise the value of ν. The most frequently used values are $\alpha = 1$ (Liu et al, 2000) and $\alpha = 3$ (Seifert et al, 2001). For this work, it was decided to use $\alpha = 1$ which was shown by Mazoyer (2016) to best represent fog droplet size distributions and also for consistency with the ICE-3 value.

From previous studies examining the value of ν where $\alpha = 1$ (Geoffroy, 2010; Miles, 2000) it was decided that a range of $\nu = 6.8$ to 11.1 should be used. The modified gamma distribution with these values is shown in figure 5. Though there may be correlations between the LWC and the value of N and ν, a parameterisation for the values of ν and N_0 for fog in the context of cloud radar has yet to be performed. For this reason, the parameters ν and N_0 are treated as varying randomly for the purpose of investigating the uncertainty in simulated reflectivity.

Comment: Please explain why (c) only has 20 events? What happened to the other 11 events?

Response: Out of the 31 events observed, 21 could be matched to a modelled event. Among the 11 missing events, 10 events could not be matched to a modelled event. For the events to be matched, fog must be present with a maximum of a 6 hour difference between the model and observation space (i.e. the dissipation time in one space can occur a maximum of 6 hours before the formation time in another). The last missing fog event was discarded due toformation time difference greater than six hours, hence it is not shown on the histogram.

Change: Added: 'Out of 31 fog events observed, 21 could be matched within the twelve hour window to a simulated event meaning that 10 observed events could not be matched to a modelled event'

Comment: If one wants to improve fog forecast, shouldn't we worry more about those 11 events? Can the authors comment if the newly selected background profiles will help improve the forecast for those 11 events?

Response: Indeed, it was for this reason that it was decided to use the MRP method to correct for spatial as well as temporal errors. However, if there is no fog forecast by the model anywhere in the extracted domain then the MRP method is not able to select background profiles which contain fog. The MRP method increases the number of background profiles containing fog for times when there is fog in the observations. A background profile containing fog was found for 73% of observation fog blocks using the MRP method, compared to 63% when the nearest grid point profile was used.

For the remaining 27% of cases where a background profile containing fog can still not be found, it may be necessary to use a climatological background profile for future retrievals.

Comment: Additionally, the caption is confusing. Do you mean "where the event "occurs/dissipates" later in the observations"? If statistics are derived using simulations minus observations, then it is best to be consistent throughout the manuscript (e.g., fig. 2 and fig.3).

Response: All statistics throughout the manuscript are for Observation -Simulation.

Change: Re-worded to 'fog formation time differences for matching events; fog dissipation time differences for matching events (differences are positive where the fog forms/dissipates later in the observation).'

Comment: Do you mean fog thickness can be exchanged with fog top height, since the figure title is fog thickness, not top height?

Response: Yes, exactly. From what I understand, the two terms are often used interchangeably (e.g. Román-Cascón et al., 2015). as long as altitude is given above ground level.

Change: The first time the term is used (section 3.1) I have indicated that the two terms can be used interchangeably, but have replaced fog top height with fog thickness where the term directly references a figure with this term.