Manuscript review amt-2020-166 « Improvement of numerical weather prediction model analysis during fog conditions through the assimilation of ground-based microwave radiometer observations: a 1D-Var study ».

First of all, the authors thank the two anonymous reviewers for their positive feedback and helpful comments to improve the manuscript. All modifications have been taken into account ; most relevant are highlighted in red in the new version. We hope this new version will be suitable for publication.

Reviewer 2 :

Major comment:
I understand the discussion about optimal estimation in section 2.3 and 3.1, however I am a little bit perplexed by the discussion of the background covariance in section 3.2.
From what I could understand the a priori vector Xb = (T, Q, LWP) used in the convergence scheme specified in line 110 is provided by 1 hr AROME forecast profiles.
The corresponding background covariance B associated with this a priori estimate was estimated as described in line 145-150 for all cases and for a subset of fog cases and the diagonal terms were multiplied by 0.7. Where I get lost is the next section (3.2) where the background covariance is modified in a seemingly arbitrary fashion by removing the cross-correlation terms from the climatology. I entirely understand using a fog covariance for the fog cases and a climatology covariance for the non-fog case. However, to "choose" the background covariance that optimizes the retrieval results seems a little bit unorthodox.
My point is that the covariance should be an "objective" way (as far as possible) to quantify the uncertainty associated with the a priori information. It seems that if the background covariance associated with the Xb is not good enough for the retrieval perhaps a different choice for the a priori Xb should be made (i.e. not from the model but perhaps from a radiosonde ensemble). Alternatively, the convergence could be controlled with a multiplicative factor to B (usually called g) that is reduced at each iteration based on the behavior of the cost function. This approach is mostly used for infrared retrievals, but, in this case, it may prove beneficial as well. So perhaps I am not entirely understanding this part, in which case this procedure of "choosing" B based on the retrieval results could be better justified or may be the straight optimal estimation approach should be modified the way mentioned above.

We would like to thank the reviewer for this very interesting question. It seems that our original paper did not justify enough the approach of possibly zeroing the cross-correlations between temperature and humidity, and this is an important point. We shall first remind that the treatment of humidity in 3Dvar has received a long attention in the NWP community (see e.g., Dee and Da Silva MWR 2003). In many operational schemes, the treatment of humidity has long been univariate (e.g. zeroing or not computing the cross-covariance between humidity and the other variables). This is for instance still the case in global ARPEGE 4DVar and in various 3DVars (eg., Barker et al MWR 2004 for the WRF model). One explanation is that the cross-covariances between humidity and temperature is flow-dependent. To be more specific, Holm et al (2002) have shown with the global IFS model that even the sign of t-q correlations was changing, being negative in average but positive in saturated conditions. Dee et al (2002) also confirmed that, if observations are not as abundant for temperature and humidity, it is more accurate to neglect mixing ratio–temperature error covariances. Whatever the used background profile (from a NWP model or from radiosounding climatology), if the analysis increment is mainly driven by the coupling, using fixed covariances for different types of weather conditions will potentially degrade the analysis. In our case, keeping cross-covariances between temperature and humidity in the climatological B-matrix is therefore suboptimal and possibly not as good as using a univariate approach for humidity, depending on the prevalence of saturated conditions. Therefore, we consider the zeroing of temperature-humidity covariances in the climatology as a natural approach, which is of course outperformed by the use of the ensemble that brings this lacking information about saturation during fog conditions. A clean way to make the B matrix flow dependent through the 6-month study, would have been to use the AROME ensemble data assimilation (EDA) to extract hourly B matrices depending on the weather conditions for each new retrieval. Unfortunately, the AROME EDA was still in development at the beginning of the study. Only a few days were post-processed and the choice has been made to select the IOP1 with interesting fog events in the observations in order to focus on the impact of an optimal B matrix during fog conditions. As this paper mainly focuses on fog retrievals, optimizing the B matrix for all other weather conditions is beyond the scope of the paper. As an alternative approach, for non fog conditions, we decided to follow the approach still used operationally in our 4D-Var scheme by zeroing the cross-correlations between temperature and humidity to avoid the degradation in the retrievals of specific humidity.

In order to clarify this aspect ,section 3.2 has been modified introducing the explanation on why different B matrices are evaluated :
As cross-covariances highly depend on the weather conditions (Hólm et al. (2002), Michel et al. (2011)) and the use of fixed covariances is not optimal when dealing with different atmospheric scenario, Config1 aims at evaluating the impact of the cross-correlations between temperature and humidity on the retrievals. To that end Config1 corresponds to the same configuration but removing the cross-correlations between temperature and specific humidity. It can be noted that this approach is still used in various 3D/4D-Var operational schemes (Barker et al. (2004)).

We also included an additional paragraph at the end of the section to clarify the conclusion :

Figure 3 confirms that the best configuration in terms of B matrix corresponds to Config2 compared to the CTRL configuration. In fact, the use of a "climatological" B matrix with cross-correlations degrades both temperature and humidity retrievals but more significantly specific humidity up to 4 km. Overall, these results confirm that, for MWRs, humidity increments in the lowest levels are significantly driven by the cross-correlations between temperature and humidity. These correlations (sign and amplitude) being highly dependent on the weather conditions, the B matrix should ideally be updated for each profile. When it is not possible, the use of a block diagonal B matrix might be preferable to avoid degradation in the retrievals due to inaccurate cross-correlations. This result is in line with the study of Dee et al 2002 which showed that, when humidity is less adequately observed than temperature, it is more accurate to neglect humidity – temperature error covariances. However, when an adapted flow-dependent B matrix is used, the specific humidity analysis is improved. In the future, the use of ensemble data assimilation schemes should enable deriving optimal B matrices evolving in time and space to be consistent with the weather conditions.

Minor comments:
Abstract: There is terminology that is not defined for example, in lines 11, 12, 15, what are 1D-var increments? I suggest either making the abstract less detailed about the results or defining the terms used.

→ We have removed some details from the abstract and we do not use the word « increment » anymore. We defined this term later in section 2.3 :
Through the manuscript, the atmospheric state minimizing the cost function is called the "analysis" (xa) and "increment" refers to the difference between the a priori xb and the analysis.

Table 4 is not clear. The caption says "Error reduction (%)" over the background. It is not clear what the negative number means. Does it mean that the retrieval is actually increasing the RMSE with respect to the background?

RMSE are always computed with respect to in-situ measurements on the tower.
Then two RMSE are computed :
RMSE_xb : errors of the background with respect to the tower
RMSE_xa : errors of the analysis with respect to the tower
What is called error reduction in the manuscript is defined by :
ER=1-RMSExa/RMSExb

when this number is negative it means that the retrievals increase the errors with respect to the tower measurements compared to the background (the 1D-Var retrieval is thus worse than the background statistically).

We modified the table 4 caption to clarify :
Reduction in the RMSE with respect to tower measurements after the 1D-Var analysis (RMSExa) compared to the background (RMSExb) for all weather conditions (upper part) or only fog events (lower part) : ER = 1− RMSExa/RMSExb (%)

Table 4: Just to make sure I understand correctly, this statistic is computed by taking one layer of the retrieval profile grid corresponding to the tower height of 50 and one layer corresponding to 120 m? I would imagine that the retrievals at these two layers are highly correlated because the vertical resolution of the radiometer at this height is about 100 m. Therefore, is there even a merit to look at two layers vs. averaging the tower and radiometer measurements between 50 and 120m?

→ We agree with the reviewer that the two retrievals are potentially highly correlated. This is now stated in the following sentence, reported at line 179 of the revised manuscript. However, we prefer to keep the comparison at both levels, as it conveys some information, e.g., that the reduction in T(K) is higher at the lower level.

It is important to note that given the relative low vertical resolution of MWR retrievals, the retrievals at 50 and 120 m are likely to be highly correlated.

Fig. 3 It is not clear what "closest GP B clim" mean in the labels.

→ We agree and have made labels consistent with the names of the different configurations defined in table1.

However later on, in Fig. 5, I see that the background configuration reappears in the right panel. Is this necessary? It has already been established that this configuration should not be used. In addition, by just looking at the middle and right panels of Fig. 5 the differences appear to be really minimal.

→ As the largest differences are limited in time (only when a correlated fog B matrix is used for humidity between 4 and 9 UTC) and vertical spread (the first 500m), it is true that the differences can appear minimal when looking globally at the figure. To avoid confusion, we decided to remove the control configuration in the right panels and show only the optimal configuration 3.

The discussion of Fig 5 is not clear. The text says at lines 240 "We can note the large temperature warming by up to 5 K from 0 to 12 UTC during the whole fog event (in the model space) in the 0-500 m vertical range." By "model space" the authors mean the left panel (i.e. forecast?). If I look at it is not clear what is meant by the large warming from 0 to 12 UTC. Is this the sharp increase in the temperature *above ~100-200 m. Is this a visual product of the rainbow color scale used? I wonder if it would be more realistic to use a continuous color scale for these plots.*

→ « In the model space » means when the model simulates fog. It only refers to the time period 0 to 12 UTC which corresponds to fog only in the model simulation. In the observation fog is only observed between 4 and 9 UTC. If we look at the temperature between 0 and 12 UTC in the first ~ 200 meters, we can see that the temperature is much colder on the left panel (the model background) compared to the middle planel (1D-Var). We just wanted to stress that this time window corresponds to the simulated fog event (and not the time where the fog was actually observed). In order to make the figure more readable, we limited the y axis to the 0-1 km range and we clarified the text :

We can note the large temperature increment, up to 5 K from 0 to 12 UTC essentially in the first 250 m, after 1D-Var is applied ; this is the period where the model simulates a thick fog event not confirmed by the observations.

Fig 6 Is there a reason why the 1D-var overestimates specific humidity between 4 and 9 UTC? Are the brightness temperatures affected?

→ The most reasonable explanation is that the positive cross-correlations between temperature and humidity are probably over-estimated. However, we are probably within the retrieval uncertainty. One sentence has been added to the manuscript :

Though closer to the in-situ observations, 1D-Var retrievals slightly overestimate specific humidity between 4 and 9 UTC. This is most likely due to over-estimated positive cross-correlations between temperature and humidity in the B matrix

Fig 7 is time UTC?
→ Correct, this has been modified in the figure

Fig. 8 and related discussion. Does the introduction of the MWR data improved the statistics of undetected and false alarm? I see that the temperature errors are reduced, but is the number of false alarms and missed detections the same?

→ This is a very good question that we investigated. However, as the 1D-Var only deals with the integrated liquid water (LWP) and does not have enough information with the MWR alone to correctly localize several cloud layers, it is complicated to interpret new statistics based only on the 1D-Var analysis. In fact, if there is already a fog or a cloud layer, the 1D-Var will normalize the vertical distribution so that the integrated path is closer to the observation. If the model initially gives no fog or cloud whereas the observation is cloudy, the 1D-Var will add a new fog layer at the level of maximum relative humidity (which may or may not be close to the ground). The scores to determine false alarms, missed events and good detections are, on the other hand, based only on the LWC at the ground.
It means that if there is a cloud layer on top of the fog, the LWP might be increased erroneously at the ground due to the cloud aloft during false alarms. If we look at the background profiles during false alarms, only 33 % do not have cloud detected by the ceilometer. It means that in the remaining 67 %, the model simulates fog at the ground though it is not detected by the visibility sensor and a cloud aloft is observed in the measurements. For these cases, the 1D-Var cannot remove the liquid water content at ground (as there is a cloud aloft) and might even increase the LWC. If we just evaluate the 1D-Var retrievals based on the false alarms scores we will conclude that there is either no impact or even a degradation. The only way to really make a conclusion would be to let the

model physics get the liquid water content profile balanced with the new temperature and humidity profiles before recomputing the score.
The same problem occurs when evaluating missed fog events if there is not liquid water in the background at the ground. However, for this specific study, most of the background profiles have a small amount of liquid water at ground level, even though the visibility at ground is not reduced to less than 1000 m. The rate of missed fog events is decreased from 27 % in the background to 19 % in the 1D-Var analysis. Again this evaluation is only based on the LWC change at the ground and it would be interesting to evaluate the impact of the new temperature and humidity fields on the LWC after a few time steps of forecasts but this is beyond the scope of this paper. This investigation into the forecast impact will be studied within the framework of the SOFOG3D experiment.

We included a discussion on this topic in the new version of the manuscript at the end of section 5 :
The next natural step of this study would be to calculate updated scores of fog detections with the new 1D-Var analyses compared to the background profiles. However, forecast scores are only based on the LWC at ground whereas the 1D-Var works on the liquid water path without information on the cloud vertical structure. During false alarms, conclusions on the impact on forecast scores are complexified by the presence of cloud layers above fog in a majority of false alarms which can cause an increase in LWC at ground. As for the hit ratio, it is increased from 73 % in the background to 81 % in the analysis. The rate of missed fog events is also decreased from 27 % in the background to 19 % in the 1D-Var analysis. However, as this evaluation is only based on the LWC change at the ground, it is necessary to evaluate the impact of the new temperature and humidity fields on the LWC after a few time steps of forecasts but this is beyond the scope of this paper. This investigation into the forecast impact will be studied in the future within the framework of the SOFOG3D experiment (section 6).

Fig. 12 x axis title is missing :
→Lables for x axis are now reported in addition to the figure caption.

Fig 13 axis labels are missing, and fonts are very small
→ We have recomputed the figure with labels and titles increased.