Comments on "Improvement of numerical weather prediction model analysis during fog conditions through the assimilation of ground-based microwave radiometer observations: a 1D-Var study" by P. Martinet et al.

Summary:
The paper examines the assimilation of ground-based microwave radiometer observations into a Numerical Weather Prediction model and concludes that radiometers can benefit the forecast of fog events by reducing the temperature errors in the model.
I found the work interesting and generally understandable, I have only a few comments on the discussion of the background covariance and a few clarifications that may be needed to make the paper more accessible to a general audience.

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 γ) 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.*

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.*

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?*

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?*

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

In section 4 line 203 is said that "*the best retrievals are with config 3 and config 4*" and later at line 225 it is said that "*Overall best performance of the 1D-Var is found with bias-correction applied on channels 1-10 and a dedicated fog B. This configuration is used in the following sections*." Taking these two statements together one can deduct that this corresponds to config 3, however it would be easier to explicitly state this at the end of section 4 (i.e. say: "configuration 3 is used in the following sections"

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.

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.*

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

Fig 7 is time UTC?

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?

Fig. 12 x axis title is missing

Fig 13 axis labels are missing, and fonts are very small