Reviewer 2

This paper details tests of the latest version of the PRO operator using global and km-scale models. This operator is compared with an operator (operadar) which was developed for the calculation of refractive effects with a polarimetric radar. The operator was also used to compare the microphysics schemes used in the global and regional models. The paper is well-written and provides interesting results. Therefore, it should be acceptable for publicaion after a few small changes are made.

The authors would like to thank the reviewer for its careful review of the manuscript and for providing valuable suggestions which significantly improve the quality of the work. In accordance with the suggestions, the authors have thoroughly checked the manuscript and shared the point by point responses to the comments. In the revised pdf file, the scientific changes are highlighted in violet. Other corrections, which are linked to some corrections made by the authors itself are displayed in blue in the revised pdf file. They are written and explained at the end of each answer.

Larger comments

L180: While the Bayesian inversion section makes reference to the paper of Paz et al (2026), I feel that this section needs to outline the strengths and weaknesses of this technique at this point. I acknowledge that there's some discussion alongside the results, but the reader would benefit from having discussion in section 2.3.1. In particular, using optimisation is effectively drawing the model's estimates of differential phase towards the observations. If this is done for a small subset of observations (as happens later in the paper) then this optimisation may have a substantial effect. This should be discussed.

The Bayesian inversion applied in this study has both strengths and limitations that shape how the output should be interpreted. Its principal strength is that it lets the scattering efficiency of each frozen species be constrained directly by the PRO observations themselves. Its principal limitation is that, because x is optimized against the very same $\Delta\Phi$ observations subsequently used to assess the operator's performance, the resulting agreement between observed and simulated $\Delta\Phi$ reflects in a part a calibration exercise rather than solely and independent validation. This effect is strongest when the number of occulations entering a given fit is small relative to the number of free parameters, as is the case for the region-specific retrievals. We treat the regional retrievals as a diagnostic and indicative rather than as independently validated parameter estimates. We have stated this aspects explicitly in Section 2.3 and in the Conclusions.

In general, the majority of conclusions made in the article would be benefitiated and strongest if the number of observations is larger. Since PRO is a relatively new technique, several instruments are starting to incoporate it, however, we still do not have enough as to make a better statistical study.

L317: It's unusual to choose the median of the Delta Phi profile as the representative level to plot. I wonder whether this affects the results, since the median value could come from different heights in the observed and simulated profiles. It would be more normal to choose a representative impact height to plot. Can the authors comment on whether the results change if the more usual choice is made?

We thank the reviewer for this comment. We have replaced the median-based scatter plot with a comparison at two different tangent heights layers, 1-2.5 km and 5-6.6 km (new Fig. 6), and repeated the correlation analysis for both the PRO operator and operadar at each height.

In the lower layer (1–2.5 km), both operators show weak correlations with the observations (PRO: r=0.22, slope =0.16; operadar: r=0.35, slope =0.22), and neither clearly outperforms the other. In the upper layer (5–6.5 km), the PRO operator follows the observations considerably better (r=0.46, slope =0.53) than operadar (r=0.34, slope =0.08). The latter simulates very low values for all occultations, whatever the observed amplitude.

The results therefore do change with this choice, as the reviewer anticipated. The correlations are weaker than those obtained with the median of the whole profile, and the ranking between operators depends on the layer: operadar has a slightly higher correlation in the lower layer, while the PRO operator performs clearly better in the upper layer, particularly in amplitude (slope 0.53 versus 0.08). The height dependence is also expected: the $\Delta\Phi$ signal peaks at different altitudes in each occultation, and the upper layer is closer to the altitude range where it typically peaks and where frozen hydrometeors, whose scattering efficiencies are retrieved from the observations, dominate. We also note that part of the PRO advantage is expected, since its frozen-hydrometeor efficiencies are calibrated against the same observations.

L416: The authors choose to adopt the parameters from AROME on the global dataset as the ones to use. However, this gives the minimum allowed weight to ice and graupel in the forward operator. At various points the authors conclude that the PRO indicates that snow is the dominant contributor to the differential phase (e.g. L441, L493). Since the optimisation is used, then this conclusion is heavily influenced by the extremely low weight given to ice and graupel. Perhaps it would be good to soften these conclusions, since they could change if a different optimisation result were achieved.

We thank the reviewer for this important observation. The reviewer is correct that, in the global AROME retrieval, x_{ice} and $x_{graupel}$ both saturate at the prescribed lower bound of $10^{-3}\mathrm{mm(kg\cdot m^{-2})^{-1}}$ (Table 2), while x_{snow} takes a finite value. This means the pooled dataset does not statistically require a contribution from ice or graupel beyond the numerical floor, and the resulting "snow-dominant" picture at the global level is partly a consequence of this saturation.

We note, however, that this is not a free-floating degeneracy of the fit: the regional retrievals (Table 2) show that x_{ice} and $x_{graupel}$ do *not* saturate everywhere, in the two most convective regions, Polynesia and New Caledonia, x_{ice} and $x_{graupel}$ are instead the dominant parameters, while x_{snow} collapses toward the lower bound. The global retrieval is numerically dominated by the more stratiform/mid-latitude cases (France, Antilles, La Réunion: 19 of 27 occultations), so the saturation of ice and graupel at the global level reflects the composition of the pooled sample rather than a universal property of the PRO/AROME system.

We agree that the manuscript currently overstates this as a general finding. We have revised the text to soften this conclusions. Also, disentangling genuine microphysical negligibility from parameter degeneracy would require a bigger amount of cases and an identifiability analysis that is beyond the scope of the present study, and we flag this as a limitation.

Smaller comments

L43, L45: Delta phi appears twice in these lines.

Corrected.

L76: Since on La Reunion is renamed, I suggest rewording the sentence to highlight just this.

Figure 1 caption: Using just "Reunion" is a departure from the naming convention.

L107: Reunion Island is a departure from the above-specified naming convention.

We have renamed all as La Réunion in order to be consistent.

L134: Upgrade is not the right word to describe the DA system's work. I suggest that "calculate" would be better choice.

Thanks for the suggestion, corrected.

L146: I believe that "strech and titled" should be "stretched and tilted".

Corrected.

L210: "graupel specie" should be "graupel species" (the word species is both singular and plural).

Corrected.

L271: I think it would be clearer if NBF is in the parantheses: "due to nonuniform beam filling (NBF, see for example..."

Corrected.

Figure 4 caption: The shaded region shows the 16th to 84th percentile range. This is an odd choice of range - why not use 10-90 or 5-95 as is normally the case?

We agree that the 10th-90th percentile range is more conventional and more easily interpreted choice than the 16th-84th range, which we had adopted as the equivalent of ±1σ for a Gaussian distribution. Figure 4 has been updated so that the shaded region now shows the 10th-90th percentile range, and the caption has been revised accordingly. The median profiles and the main conclusions drawn from the figure are unchanged.

L473, L476: I'm unsure what is meant by the text "5,km". I assume that the comma is not indended?

Yes, the comma is not needed. Corrected.