Response to Anonymous Reviewer 1

The authors would like to acknowledge the reviewer for his comments, which gave us the opportunity to clarify and improve the paper.

Point to point response:

General comments:
In the abstract and the introduction you should mention the calibration using rain gauges, because as far as I know it's one of the most used methods for the calibration (quantitatively).

We will mention the use of rain gauges for "calibration" purposes in the introduction. However, it is our opinion that use of rain gauges should be intended more properly for validation, not for calibration. As such, it was not considered in this work.

In the paper, there are a lot of repetitions:
-In the abstract L5-6 and L16-17
-Page 6, L5-7
-Page 8, L26-30
-Page 10, L25-26

Repetitions deleted.

The inter-calibration technique: it is not a new development; it was already developed in France and in Canada, please refers to:
-Zlatko Vukovic, Environment Canada, Canada, and M.C. Jim Young, Norman Donaldson: Inter-radar comparison accounting for partially overlapping volumes, Erad 2014, Garmish Partenkirchen Germany.
-Ribaud, J.-F., Bousquet, O., Coquillat, S., Al-Sakka, H., Lambert, D., Ducrocq, V. and Fontaine, E. (2015), Evaluation and application of hydrometeor classification algorithm outputs inferred from multi-frequency dual-polarimetric radar observations collected during HyMeX. Q.J.R. Meteorol. Soc. doi:10.1002/qj.2589
These references have the same method, perhaps it is not 100% similar, but for me it is enough similar and they should be mentioned!

Thank you for providing these useful references, which we have now included in the paper.

<u>Specific comments: (P: page, L: line)</u>
P2, L29: why only the first and the last day?

We have received similar comments also from the others reviewers, so we decided to extend the self-consistency technique to all the significant rainy events, i.e. the same events that are used for the inter-calibration.

P2, L29: What is "proper meteorological conditions" means, can the authors specify the characterizations.

For this study, we subjectively chose the cases according to the daily rain gauges analysis. As future development, we will automate this method using proper thresholds on the number of radar observations in precipitation events.
Since other reviewers suggested to expand the self-consistency procedure to the whole study period, we decided to perform this technique when precipitation occurs in the radar domain.

Pa3,L5:Please add (asl) so you can use it correctly later, i.e L16 of the same page.

Added, thanks.

P3, L11: ZH and ZDR are already defined.

Removed "radar reflectivity" and "differential reflectivity".

P3, L12: 50 pulses for H and V or 50 pulses for each?

50 pulses for each polarization.

P5, L14: "same frequency": just as remark, I believe that such method is much more valuable when you use 2 different wavelengths to better study the PIA, PIDA and for microphysics!

This is true, but for calibration purposes it is actually better to use different radars operating at the same frequency, to avoid having to consider different scattering regimes.

P5, L15: theoretical model: what you mean by that, normal propagation?

Yes, normal propagation (atmospheric standard conditions) of the radar beam

P5, L16: sufficient met echoes, how much?

In terms of reflectivity pairs, "sufficient" means more than 100 pairs, in order to allow computation of significant statistics. It is now clarified in the text.

P5, L19: I understand that the radar cells should have similar size, but what is your percentage of tolerance.

We are not using percentage because the actual value would depend on the location of the target volume. We use fixed values of tolerances: the vertical tolerance is 100m and the difference of the distances from the two radars has to be less than 40km.

P6, L11: 10^5 is by time step comparison (every 10 min)?

Yes, it is for a given time volume scan comparison. The volume scan is repeated every 10 mins, so the theoretical maximum number of pairs during a day would be in the order of 10^7.

P6, L21: perhaps you should mention here the abnormal propagation (anaprop)?

Yes, we now discuss the anaprop in the ground clutter calibration technique. In Fig. 18, we can note that the ground clutter calibration shows an increase of the 95 percentile for the Bric della Croce radar around 26 September 2014. Re-analyzing the data and using the nearby radio sounding observations to compute the refractive index of air, we can conclude that in those day the anomalous propagation influenced the ground clutter calibration, bending the radar beam toward the ground. In fact, the 95th percentile of Zh increased by about 1dB.

P7, L1: here the authors are talking about all the ground clutter or the nearest one? Is there any threshold on ZH?

We consider all the ground clutter echoes. There is no Zh threshold for the clutter mask generation, only the frequency of occurrence matters.

P7, L10: and 3d, a space is missing.

Corrected.

P9, L3: this part is tricky, the authors are using a product to evaluate a moment, BUT this moment is used to calculate the product!

We understand this doubt of the reviewer. However, we believe that we can exclude this possible effect considering that:

- the hydrometeor classification relies on several variables, not only Zdr, and for the liquid/solid discrimination in general the upper air temperature observations are especially relevant
- In particular the algorithm employed in this work (Bechini and Chandrasekar, 2015), was shown to be relatively insensitive to small biases in Zdr

Bechini, R., and V. Chandrasekar, 2015: A Semisupervised Robust Hydrometeor Classification Method for Dual-Polarization Radar Applications. Journal of Atmospheric and Oceanic Technology, 32, 22-47

P9, L6: why this interval for ZH?

This interval should correspond to drizzle or light rain conditions, with a distribution of nearly spherical drops (the expected Zdr is close to 0 dB). The lower limit (10 dBZ) is used to avoid noise-contaminated observations.

P10, L7:-0.49 basing on the figure and not -0.5

Corrected.

P10, L21: please remove s from models; I guess the authors are using one NWP model.

This is correct, we removed the “s”.

P10, L25: Are you taking a ring around the radar or some specific points.

The intercalibration in performed on the range-bin pairs computed from the beam propagation model with the aforementioned thresholds, as reported in Sec.3.2.