Review of: "Calibration of radar differential reflectivity using quasi-vertical profiles", by Daniel Sanchez-Rivas and Miguel A. Rico-Ramirez

General Comments:

This manuscript offers a method for monitoring the ZDR offset of a dual-polarization radar using quasi-vertical profiles (QVP). The method is applied on C-band weather radars in light rain media. The authors suggest accuracy to O[0.1 dB], e.g., potentially in-line with 'bird-bath' calibration (natural media). There are two apparent justifications for this publication: its improvement compared to previous natural media efforts, and its QVP application towards these ideas.

The manuscript is not recommended for publication. The study is functional with elements similar to the typical AMT scope, but the reviewer finds low value in the 'new' concept/application. The use of intrinsic liquid properties for ZDR monitoring is well known, origins in low angles and selective ZDR averages (i.e., cell peripheries). This manuscript adds a "QVP"-wrapper aimed now at liquid media, yet lacks the physical underpinning as to why such methods would improve performance over a boilerplate practice to 'average ZDR in light rain'. These "QVP" concepts are evaluated against a modest dataset, but reads to the reviewer as motivated by convenience and applying a poorly-matched "QVP" concept (hammer looking for a nail?) in a less-behaved condition (light rain) to be 'novel'. Yet, it seems a straightforward evaluation of an existing snow QVP application (as less original as that seems) may have been far less controversial. The authors perhaps unintentionally increased their degree of difficulty (at least, to this reviewer), by leaving the reviewer questioning whether simpler, quicker, or (existing dry snow) options for targeted averaging may be equally/more effective.

Moreover, a central claim for this effort seems to follow its 'relative' calibration performance (oversold), esp. for "light rain". It is unlikely any 'natural' method can genuinely guarantee accuracy better than 0.2-0.3 dB – this has been well-argued by previous authors, including several cited; Prior efforts were rightfully cautious in their claims. Yes, some allowance can also be extended to older studies that are occasionally captives to their moment (i.e., radar technology improves with time → better ability to target lighter rain, etc.). Nevertheless, the intrinsic "light rain" variability is significant and comes in many forms (not limited to):

- Capabilities to provide 'ground truth' (e.g., disdrometers as a poor light rain reference);
- What gets defined as 'light rain' (regional / physical process variability),
- How one identifies these regions with existing radar (Z calibration, etc.), and
- Location, radar sensitivity/quality, other vertical profile factors (e.g., evaporation, sorting, process) that undermine accuracy claims when averaging over regions.

For this reviewer, the authors have not demonstrated they built a better mousetrap. The reviewer understands there is an inevitable overconfidence (aka, marketing) in most manuscripts. However, "relative", not absolute calibration concepts are typically quite conservative, and it should be obvious that selective performance may be better under ideal conditions. The authors' disdrometer image (Figure 5) alludes to some inherent variability in (surface, 'instantaneous') ZDR properties in "light rain" (aka, dynamic range of intrinsic ZDR > 0.6 dB). These depictions are consistent with discussions by Bechini et al., Ryzhkov et al., for what those authors expect from "light rain", or why "light rain" (generic) is less suitable than "dry snow" (**see also, specific comment**). Select locations (UK) experience different bulk microphysical expectations (e.g., propensity for widespread rainfall, stratocumulus), thus performances may reflect strong local process / natural advantages (e.g., contrast with "light rain" at the peripheries of thunderstorms).

Overall, one takeaway message is that this reviewer does not feel the authors have justified the "QVP" application as a genuine improvement over a generic "average" ZDR monitoring practice, for rain, snow or otherwise. Rather, the reviewer claim may be that "QVPs" in light rain are arguably *far worse*, given this form of averaging enables mixtures of less suitable profile properties that produce apparently viable "light rain" profiles. Why use a "QVP" process at all? Fundamentally, this is a reduction of information; Many previous studies speak to physical 'profile' issues convolved with "QVPs" and similar averaging, with even the QVP originators shifting to "CVPs" or other targeted averages – For example, ZDR should naturally evolve below the melting layer in response to processes such as sorting, evaporation, break-up, and/or other regime-averaging nuances (within event, or tropical vs midlatitude differences). This all points to why previous studies may have remained cautious in their claims on relative 'light rain' use and uncertainty, but also where QVP-ideas are suboptimal (esp. in rain, below cloud, etc.). The reviewer is questioning the need in using a QVP in these contexts if the QVP cannot be justified as out-performing any number of simpler, targeted ZDR averages of 'light rain' (if one is already thresholding regions loosely on Z, RHV regardless, you've already opened that echo classification bag once one introduced decision-tree thresholding for 'drizzle', etc).

Specific Comment:

Why do the authors use "light rain" for the "QVP"? Many efforts point to why they avoid light rain (see, Ryzhkov et al, discussions). Unfortunately, the reviewer might have been more amenable to an AMT manuscript that was simply a long-term validation for an existing 'dry snow' QVP concept. That is because most "QVP" concepts and ZDR calibration at higher tilts focus on the properties of lower density, dry aggregate snow as a claimed better-case media. They often note that the spatiotemporal averaging/variability is still a concern, but perhaps less in-cloud and widespread stratiform selective events. Overall, those rationale (e.g., Ryzhkov et al. and subsequent) reflect a somewhat different take on the role of higher tilts and the expected ranges for ZDR media at higher tilts. The current authors use expressions such as:

b. The intrinsic value of Z_{DR} for angles below 90° and collected in light rain is different from zero. Also, it is elevation-dependent, as demonstrated by Bringi and Chandrasekar (2001) and formulated by Ryzhkov et al. (2005a) as:

$$Z_{dr}(\theta) \approx \frac{Z_{dr}(0)}{\left[Z_{dr}^{1/2}(0)\sin^2\theta + \cos^2\theta\right]^2} \tag{8}$$

The reason Ryzhkov et al. give for higher tilts and dry snow is seemingly opposite to the current authors' logic – Ryzhkov argues dry snow has lower natural ZDR variability, and when these media are viewed from higher tilts (e.g., the eventual multiplier on ZDR in equation (9) would be closer to 0 instead of 1), the dynamic range of potential ZDR variability is low. When the underlying media experiences a wider range of variability, aka, light rain ranges from 0.1 dB to 0.6+ dB at typical trusted Z ranges, etc., this implies added uncertainty for any 'average' reference frame. These issues are at their most problematic at grazing angles, and possibly not preferable at lower altitudes (given evaporation, other profile physical processing that evolves ZDR below cloud). Thus, it is not immediately preferable (for their concepts) to have:

$$Z_{dr}(\theta = 10^\circ) \approx 0.968 Z_{dr}(\theta = 0^\circ) \text{ [dB]} \quad (9)$$

e.g., a high coefficient close to 1 is 'bad' for "light rain" in these contexts, b/c the intrinsic ZDR for Z ~ 15-20 dBz remains in those ranges from 0.1 to 0.6 dBz (aka, author Figure 5); This drives the potential uncertainty against the 'reference' ZDR, and one may be correcting by >0.3 dB quite often (perhaps this was worse in Oklahoma, where lower, unregulated use of Z caries a wider range of ZDR). High tilt intrinsic property sampling (if available, aka, 'birdbath' at its limit) acts to limit that range of possible ZDR → better chance to accurately pinpoint ZDR. Thus, the authors' statement,

"Hence, ZDR radar measurements collected at elevation angles below 10° are similar to those collected at lower elevation angles and so they do not add additional uncertainty to the offset correction method."

… feels opposite this 'dry snow' rationale. This seems to be a question of whether the authors genuinely believe they can target low-ish variability ZDR 'drizzle' better than low-variability dry snow? This may be regional to the UK, e.g., stratocumulus w/drizzle, but may not seem as reasonable if painted with a US NEXRAD radar lens, as a separate example.

Even with light rain, these issues are likely worse than presented; For example, this effort has not fully discussed that the disdrometer (Parsivels, etc.) references are poor in light rain R < 1-3 mm/hr. It is unlikely most units capture light rain properties perfectly, esp. with assumptions made for disdrometer processing (a different subset of literature on Parsivel, 2DVD and other light rain comparisons). Dry snow media, similarly, has its own issues with identification, wavelength dependency, complications to "QVP" profiles from non-uniform beam-filling (at C-band, there is potentially intrinsic negative ZDR above the ML owing to non-uniform beam filling!). There is not a quick fix, unfortunately.