

Response to reviewer 1

This manuscript documents the first multi-Doppler 3-D wind dataset over the United Kingdom, built upon the PyDDA variational framework with several important methodological upgrades: the new ORIGAMI model-guided velocity dealiasing algorithm, multi-radar composite VAD initialisation, and a cloud-top upper boundary condition to suppress spurious vertical velocities above cloud layers. The 1-km horizontal/500-m vertical, 10-min-cycled three-dimensional wind product derived from five radars during the WesCon-WOEST field campaign represents a valuable new observational resource for evaluating convection-permitting numerical weather prediction (NWP) models. The work includes comprehensive validation against radar wind profilers (RWP), FAAM aircraft in-cloud measurements, and qualitative comparison against high-resolution CAMRa RHI retrievals.

Overall, the study is well-motivated and fits well within the scope of *Atmospheric Measurement Techniques*, as it describes both algorithm modifications and a complete, campaign-scale observational dataset intended for public archiving. The manuscript is generally clearly structured; however, several key methodological descriptions, validation interpretations, and dataset documentation details require substantial improvement before acceptance. Therefore, a major revision is therefore warranted.

We thank the reviewer for this summary and their comments and suggestions. We have updated the manuscript in response, as well as performing some additional analysis. We believe that the paper has benefited from the clarifications and further discussion added in response to both of the reviews.

Major comments

1. The ORIGAMI algorithm is a core novelty of this work, using NWP-simulated radial velocities to select the optimal Nyquist fold for observed radar radial velocities. While the four-step workflow is summarised, critical implementation details remain absent. For example: (1) What spatial/temporal interpolation is applied to map WRF model winds onto each radar's polar coordinate system? (2) What tolerance or threshold is used to flag cases where model winds are too discrepant from observations for reliable dealiasing? The paper mentions misplacement of intense convective cells in WRF can degrade ORIGAMI performance, yet no quantitative diagnostic is provided to identify such problematic volumes in the final dataset.

We accept that complete implementation was not included. We have included an additional paragraph more details on the implementation in terms of timing and spatial matching, this is however a simple case of nearest available forecast and pixel. Exact implementation in other situations would likely be case specific, however we appreciate including our precise setup informs the reader.

We note that the required accuracy is not difficult to achieve, and instances where dramatic changes occur (particularly winter storm fronts) are not within the data or focus of this study, if such instances were important to the use then an allowance for timing would be required. We have seen only a single instance where the unfolding is incorrect in our dataset. We have

however added to the conclusions that a study comparing the reliability of the ORIGAMI algorithm, relative to alternatives would be valuable, particularly in potentially more challenging conditions.

2. The manuscript correctly identifies spatial-temporal scale mismatch as the primary driver for PyDDA's underestimation of peak updrafts and downdrafts compared with FAAM aircraft and CAMRa RHI products. However, the discussion conflates three distinct error sources: intrinsic smoothing from the $1\text{ km} \times 1\text{ km} \times 500\text{ m}$ grid resolution, temporal smoothing from the multi-minute radar sampling window, and limitations within the variational cost-function setup. At present, it is difficult for readers to disentangle how much of the vertical-wind bias arises from each factor. I strongly suggest the authors explicitly separate these sources of uncertainty in Section 7: Discussion. Where possible, add quantitative estimates: for instance, comment on what horizontal scale of convective updrafts can realistically be resolved given the radar sampling geometry and chosen grid resolution. Explicitly state these resolution-derived limits within the dataset metadata description so future users can appropriately interpret updraft statistics computed from this product.

We thank the reviewer for this comment and suggestion as we agree that this is important. We have updated the discussion section to separate these sources of error and discuss them individually, in so adding sections on “Temporal Smoothing from the Sampling Window” and “Limitations from the Variational Setup”. While it is not possible to disentangle these sources of error quantitatively, their discussion and likely consequences (most significant on the vertical wind component) is certainly important.

We are currently running our algorithm for a case study using higher resolution (which would not be viable for the entire summer, but is possible for short periods) with the intent to repeat the RWP verification using the higher resolution grid, this can then be compared with the sensitivity tests on cost function weights presented in section 5. This would give an quantitative indication of how much error reduced with a finer grid, though would not completely disentangle the error sources. Initial tests suggest that the higher resolution does perform as expected, with higher peak updrafts, but maintaining the overall up and downdraft pattern. Once the results are in the manuscript will be updated to help understand the effect resolution has on the derived updrafts.

3. The manuscript presents RWP-derived wind profiles extending down toward the surface in Figure 4b, with no explicit indication of the instrument's low-altitude blind-zone. For typical L-band (1290 MHz) boundary-layer radar wind profilers, the hardware blind zone / minimum usable height commonly lies in the range of 300–500 m AGL, as documented in Guo et al. (2023, <https://doi.org/10.1002/qj.4474>). In this work, the deployed PCL1300 RWPs were operated in low-mode with the **first range-gate centred at 75 m AGL** (gate spacing 75 m). However, the lowest gates are frequently contaminated by ground clutter, transmitter leakage and low signal-to-noise ratio, so valid meteorological winds are not always reliable at these very lowest levels even if raw gates exist. This important limitation is not visualised or annotated in Figure 4b. This is physically unrealistic and may mislead readers.

We thank the reviewer for raising this important point regarding near-surface RWP observations and low-altitude clutter/recovery limitations.

We agree that for standard or unshielded boundary layer wind profilers (such as the PCL1300 systems referenced in Guo et al., 2023), ground clutter, transmitter leakage, and low SNR frequently compromise the lowest range gates. However, the RWP deployment in our study incorporated specific physical and algorithmic mitigations to ensure data validity at these lower levels. These RWP uses a Degreanes 3 peak algorithm. It identifies 3 peaks at each altitude and then performs temporal and vertical continuity checks to make sure the best peak is identified. When required 5 peaks can be identified. Other systems can and do have more problems low down.

Our RWP were both operating in custom made specialist clutter screens/ Fences/ compounds. These were designed from ¼ inch welded mesh and significantly reduced the influence from ground clutter that you can get in some RWP. The sidelobes cause the ground clutter and are blocked by the clutter screen. This means that little clutter is seen by the wind profilers and the software is good at rejecting clutter.

Comparing the wind profilers used in this study to radiosondes launched during the WesCON/WOEST campaign, these show RMSE less than 2ms^{-1} and correlations of approximately 0.94 with bias in each component smaller than 0.4 ms^{-1} .

We have updated section 2.3 (Data – Radar wind profilers) to now include detail of the clutter mitigation fencing and the algorithm which should mitigate the reviewer's concerns.

The authors can revise the main-text description to clarify: “While the PCL1300 RWP samples starting at 75 m AGL in low-mode, the lowest ~several hundred metres often suffer from ground-clutter contamination and low SNR; therefore, only data passing strict SNR quality flags are treated as reliable for comparison, consistent with typical L-band RWP performance described in Guo et al. (2023).”

Besides, please check other figures (e.g. Figure 6 time-height cross-sections) to ensure that the RWP low-altitude quality-control thresholds are either plotted as masked regions or clearly described in captions.

The 3D wind product rarely has radar observations in the lowest pixel that would be affected and the RWP used in this study have significant mitigation for the negative effects of the signal and ground clutter, given the additional detail added to section 2.3 on the mitigations, we do not believe that this is required.

4. The sensitivity tests across 16 combinations of c_m and c_o support the final choice $c_m=4096$, $c_o=0.01$. The bootstrap significance testing shows metrics are relatively insensitive across a broad parameter space when ≥ 3 radars observe a grid cell. However, the manuscript provides little practical guidance for other researchers who wish to apply this modified PyDDA workflow to other radar networks. It remains unclear whether these weight values are transferable to other campaigns or are tuned specifically for the WesCon-WOEST radar configuration.

We believe that these settings are appropriate for our use, but cannot state if the same would be appropriate for other setups and conditions, the chosen settings will however be a useful starting point for others wishing to perform similar analysis elsewhere.

We anticipate that the weighting of the obs term would depend on network density and radar data quality. The smoothing would be grid resolution dependent.

We do however believe our demonstration that the derived 3D wind grid is not particularly sensitive to the weights is likely to apply to other cases. The text has been updated at the end of the section on sensitivity to this end:

“The selected weightings for the cost function are appropriate for use with the radar setup present during the WesCon-WOEST campaign, but it is not possible to determine if the same values would be appropriate for other setups and conditions. The chosen settings may however be a useful reference and the demonstration that the derived 3D wind grid is not particularly sensitive to the weights is likely to apply to other cases.”

5. Reduced retrieval quality below 2 km is attributed jointly to surface $w=0$ boundary condition, mass-continuity constraints, and poor multi-radar beam coverage close to the surface. However, the relative contribution of observational geometry versus algorithm constraints is not quantified.

It is not possible to separate the 2 error sources, however a additional panel to figure 1 shows the low altitude radar coverage which demonstrates the extent of that problem.

Minor comments

Typographical and notation issues

- Page 9: “ORAGAMI” in Figure 3c,f caption should be corrected to “ORIGAMI”.

This typographical error has been corrected.

- In Table 1 caption: clarify what “RHIs were determined in real-time using a cell-tracking algorithm” means for CAMRa scan timing.

Updated caption to add “meaning that RHI scheduling is not a repeating pattern”. The key is this radar is not used in the 3D winds product due to irregular scanning, however a small amount of data (2 scans) are used later in the discussion section so we feel should be included in the table.

Figure improvements

- Figure 7: add short annotation in panel titles to remind readers that RWP cannot provide reliable vertical-wind for validation, only horizontal winds are compared.

Figure updated with this note in the final panel.

- Figure 13: add brief annotation explaining that CAMRa-derived w is not independent ground truth but a mass-flux-based estimate following Nicol et al. (2015).

The caption has been added to with an extra sentence: “The CAMRa-derived w should not be considered as a ground truth, but a mass-flux-based estimate (Nicol et al, (2015).”

Text clarity

- Section 3.3.2 (composite VAD initialisation): clarify how data from the five radars are weighted when fitting the sine wave for multi-radar VAD wind profiles. The text states “equally weighted”, this should be emphasised in the main body, not only in-text.

Text has been changed to hopefully make clearer that each radar is considered of equal importance in this VAD calculation.

- Section 4.2: when describing FAAM aircraft temporal-spatial mismatch, explicitly state the 5-minute window over which radar volumes are aggregated for each 10-min output product.

This PyDDA “heartbeat” was covered earlier in the paper, with more information provided in response to the reviews but it has been briefly repeated in this section for clarity “(there is a grid every 10 minutes, but that the data for that grid is collected over 5 minutes)” has been added.

- Section 7.1: when discussing CAMRa qualitative comparison, clarify that CAMRa RHI data were *not* ingested into the multi-Doppler retrieval, only used for post-hoc evaluation. This point can be easily missed by casual readers

We thank the reviewer for noting this possible confusion, we have updated the text of section 7.1 to clarify that it is not used for 3D wind generation, and explain why (generally not available and variable scan strategy).

Terminology

- When describing the “cone of silence”, briefly define this term for readers less familiar with radar meteorology.

“Cone of silence” was used twice in the initial submission, changes to figure 1 that came as a response of the reviews have removed the first of these instance, the second is now explained with a note after use “(the volume directly above a radar not sampled by PPI scans at elevations below the vertical)”.

Citation: <https://doi.org/10.5194/egusphere-2026-4631-RC1>