

Response to reviewer 2

Review of egusphere-2026-4631: “Derivation of the multi-Doppler 3D wind field during the 2023 WesCon-WOEST campaign in southern England” by Thompson et al.

Recommendation: Accept pending Major Revision

Summary: I support the eventual publication of this work given the novelty of the 3D wind dataset, but I’m concerned that the w retrievals are too erroneous (or dubious) for many applications. I think this limitation of the dataset should be plainly acknowledged, or that additional evidence be produced demonstrating the accuracy of the w retrievals, which would likely require repeating the multi-Doppler retrievals with more sophisticated methods.

We agree that the limitations of the data, particularly arising from the resolution of the derived grid, must be considered to ensure appropriate use. The discussion section (7) now includes additional subsections on sources of error in the derivation and include that the w component is particularly sensitive to the observation geometry. However, as long as the resolution and uncertainties of the data are considered we believe that this data can provide important information for numerical models, including high resolution research grade models. To ensure this is clear we have added in the penultimate paragraph which recommends model evaluation “Any evaluation would need to carefully consider the limitations of this dataset as well as the model. This would particularly need to consider timing and resolution on which the data and model are representative.”

Major Comments

The primary motivation for this multi-Doppler retrieval dataset is “to provide physically-plausible estimates of updraft strength and sizes to constrain sub-km NWP simulations”. I’m not convinced the accuracy of the w retrievals meets this standard. Based on the results shown, w retrievals seem to rarely exceed magnitudes of 5 m s^{-1} , which is not very large for the 1-km analysis grid spacing. In the comparison with flight data (Fig. 8), the retrieved w entirely fail to capture the stronger drafts. I agree that the spatiotemporal mismatch between the retrievals and flight data likely explains part of the discrepancy, but in any case, there’s little evidence here or elsewhere in the paper that the w is retrieved with enough fidelity to help guide the development of sub-km NWP simulations. Even “plausible” estimates of updraft strength and size (L432–433) would not constitute an independent (i.e., empirical) benchmark for simulated

We certainly agree that any model evaluation using observational data needs to consider the uncertainties and abilities of the observation. In the case of the $1 \times 1 \times 0.5 \text{ km}$ resolution wind field derived in this paper we should not expect the vertical velocities to reach the magnitudes seen by the aircraft, with a much smaller effective sampling volume. We do however believe that the data provides important detail beyond purely peak updraft intensity (which it cannot be expected to estimate), but also its extent, location and persistence as well as for downward motions. As an example, we refer to the work of Nicol et al. (2015) (cited in the original manuscript) that demonstrates how an imperfect vertical velocity retrieval can provide useful insights into model biases of updraft strength and size.

I believe the weak w retrievals largely arise from lacking low-level radial velocity coverage (implied by relatively high radar base elevation angles of 1.0 or 1.5 degrees). As demonstrated in

Shapiro et al. (2009) and Potvin et al. (2012), a vorticity constraint can substantially improve w retrievals in this scenario. The paper notes that the “vorticity constraint in PyDDA requires a single system motion, which is not straightforward to derive for the size of our domain and in the context of scattered convection, and is therefore not considered here.” I expect that imposing the vorticity constraint, even with locally erroneous domain-constant pattern translation velocity estimates, would have substantially improved the w retrievals. Neglecting this constraint substantially reduces the value of what could otherwise be a highly impactful multi-Doppler retrieval dataset.

As a result of this comment we have tested the vorticity constraint, using the mean flow from the VAD at 2km altitude (an approximate steering wind) as the system motion. Unfortunately we found that the vertical winds retrieved are unchanged for low weights of C_v , of up to 10^5 , but by 10^6 this term becomes a dominant factor approximately halving all vertical velocities (rather than the envisaged increase), but maintaining the structure of up and down draft locations. We have investigated further why this is happening and believe the problem is the timing of the radar data.

The impact on horizontal winds was minimal meaning that a useful sensitivity test on validation data is not possible.

This additional analysis has been discussed in the latest version of the manuscript. The term is mentioned in the methodology and in the sensitivity section a new subsection has been included:

“The PyDDA algorithm includes the possibility of including a vertical vorticity constraint in the form of an additional term of the cost function. This term requires a system motion, which is hard to define over a large domain. To test the relevance of this term and its potential inclusion, we have attempted an implementation where the system motion is determined to be that of the initialisation VAD at 2km altitude (this is approximately the steering flow for the case tested). We discovered that the output was highly sensitive to the weighting applied to this term. Once the weighting reaches a magnitude large enough to have an impact its effect is to approximately halve all vertical velocities, which is not supported by the aircraft comparisons.

The likely reason for the failure of this term to improve the retrieval is timing: Gebauer et al. (2022) suggested that a timing window of 30 s is required to avoid large errors. They showed that accounting for advection could resolve this problem to a longer collection window, however advection is not currently incorporated within PyDDA. As we cannot appropriately implement this term, and it has shown high sensitivity and a negative influence to our results, this term is not included in our retrieval.”

I’m also concerned about the lack of advection correction given the long (10-min) radar volume periods. Failing to account for the nonsimultaneity of radar observations will inevitably produce severe vertical misalignments of wind field features in many cases, substantially degrading all three retrieved wind components. Even a crude, domain-constant pattern translational velocity estimate is much better than assuming stationarity.

The reviewer has a very fair concern here, although we note that the radar volume scanning periods are actually only 5 minutes rather than 10 minutes. We have clarified this in the methodology section “The 5 radar volumes are collected every 10 minutes from the nearest

available times. Each radar's data volumes take less than 5 minutes to be collected and are timed such that they are occurring simultaneously, this means that the radar observations have a time window of 5 minutes, with a new grid every 10 minutes.”

However that still leaves timing gap in our data of up to 5 minutes, which is an inevitability from the time it takes to complete the radar volume. As in the response to the previous comment, we believe this is why the vorticity tendency cost function term is not useful.

In response to this (and the other review) we have added a section to the discussion on the consequences of the timing of radar data. Importantly we note that including advection into a future version of PyDDA would be expected to lead to improved results.

I recognize that repeating all of the multi-Doppler retrievals may not be practical in the short term. If the authors choose not to repeat the retrievals at this time, I believe the limitations of the retrieval methodology and the dataset should be clearly acknowledged in the paper.

As demonstrated above, we trialled reprocessing with the vorticity term included, which gave undesirable results, though this appears to be radar sampling issue. We are currently reprocessing a case study with higher resolution to answer questions brought up by the reviews, detail of which will be included in the manuscript.

We have made a number of manuscript improvements to ensure the limitations are fairly covered by our paper. In response to both reviews there are a number of clarifications or additions to the paper. We have also added 2 sections to the discussion section to cover in more detail the uncertainties in our methodology, now better covering the timing errors and variational retrieval setup limitations. While giving quantified uncertainties resulting from these errors is not possible, the acknowledgement of them, and the causes should help others make fair use of the data.

Minor Comments

Fig. 1: Please also show multi-Doppler coverage nearer the ground. This would provide important context for the quality of retrievals that can be achieved.

Figure 1 has been updated to use the complete geometry and an extra panel included for 750m altitude. The 250m altitude is not informative as there is very little coverage so low down. The plots are less “clean” than the original version but more informative to the realities of coverage and the added low altitude panel certainly clarifies later discussion within the paper.

In response to these changes the caption of figure 1 is updated and the body text has been changed, removing the reference to the missed cone of silence. That paragraph now reads: “Figure 1's centre and right panels show where in the region dual- or multi-Doppler wind retrievals are possible based on an angle between 30° – 150° , to give angles suitably different for deriving 2 components. These figures account for the beam elevations and spacing. Low in the atmosphere the coverage is worse as the effective maximum range of the radars to see low down becomes a significant limitation.

Combined, the five radars provide excellent coverage of the target region in the core of convective cells and weather systems, particularly in the area between 3--2°W where most of the research flights took place.”

Fig 1: Extraneous heading on panel 1.

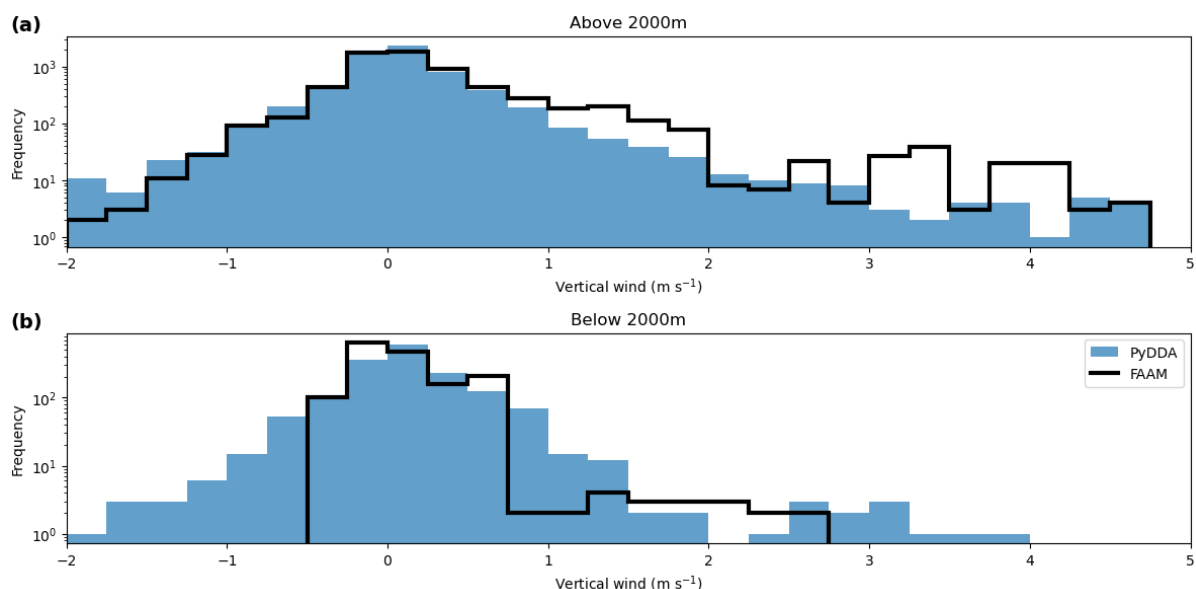
This erroneous title has been removed.

Fig. 7: ‘3D wind’ y-labels are incorrect

Changed this axis to “Multi-Doppler wind (m/s) for clarity

Fig. 9: A log y-axis would be more appropriate since some of the most important differences are in the histogram tails.

Quantisation becomes an issue with this plot with a logarithmic axis, making interpretation harder as can be seen in the figure below. The above 2000m histogram shows the PyDDA method has slightly less of the strongest updrafts, below 2000m this is less certain, though quantisation and representativity are both concerns. There is a likely a factor of representivity similar to a radar-gauge comparison: while averaging has helped extend the aircraft’s point measurement to a line, it remains essentially 1 dimensional, compared to a 3D volume defined by the multi-Doppler grid. We believe that the linear version is a fairer assessment of the true data and it’s representativity.



Line 258: windtol undefined

We also spotted that this was incorrectly formatting in the text, we have improved and clarified this sentence “The algorithm is run with a wind_tol=0.5 (the wind tolerance threshold for early stopping the optimisation) and max_iterations=25 (the iteration limit if wind_tol is not reached). This setup potentially runs the cost function >250 times, though has generally converges before reaching this limit.”

L312–313: Remember to add the data DOI!

With the potential changes to the data as part of this manuscript review the DOI is not yet set. However, the process of its long term storage and DOI assignment has been initiated.

Fig. 6: Errant colorbar label on (c)

Corrected to “number of radars”

L348–349: “reduced observation of low wind speeds” → “underestimation of low-level wind speeds”?

L349: “addition of local observations rapidly improves the statistics” – please expand on this

In fact these lines appear to have been erroneous and now corrected and clarified as such:

“Fundamental features of the design and observational geometry are the likely cause of the increased observation of low wind speeds (most notable in the u component). Note that addition of more radars providing local observations rapidly improves the statistics, with no sign of this low wind artifact once 3 radars are observing.”

Fig. 12: Colorbar for w is incorrect.

This error has been corrected in the latest version.

Wording / grammar / typos There are too many wording/grammar errors to document all of them, and they can anyway be detected with any grammar tool. The most recurring issue is run-on sentences where a comma is used in place of a semicolon or the beginning of a new sentence.

We thank the reviewer for highlighting concerns regarding grammar and sentence structure. The original manuscript was carefully proofread by all authors. During the revisions a number of minor grammatical issues and a small number of overly long sentences have been corrected throughout. We have also passed the manuscript through an automated grammar-checking tool. We have searched for long sentences and separated them where appropriate. We hope that together these changes will improve readability.