Review "High Resolution 3D Winds Derived from a Newly Developed WISSDOM Synthesis Scheme using Multiple Doppler Lidars and Observations"

Thank you for taking into consideration the suggestions made in the first round of revisions. The study benefits from the additional experiment and the added information clarifies the WISSDOM procedure and the parameter choices. Nonetheless there are still a few points that remain unclear after the revisions.

General remarks:

Overall, the level of English has worsened in the new text additions. Some sections are highlighted in the specific remarks below, but overall a language revision by a native speaker is strongly advised.

The additions in the algorithm section are very helpful, however a few questions remain. You mention using soundings for the background and as a separate constraint, how does J2 differ from J6 in the modified WISSDOM? Why do you not include all possible sources in the background? Please also mention the use of gradient descent to converge to a solution explicitly in the text.

The procedure on how to retrieve the LIDAR QVPs is still not clear to me. Could you provide a small paragraph describing the method? Do you use a VAD technique to estimate the vertical profile of horizontal wind, as suggested in Ryzhkov 2016? And if yes, how do you estimate vertical wind without having a vertical scan? Do you derive a QVP over each LIDAR or just the DGW site as suggested in Fig. 10? Please expand the text accordingly.

In Fig. 16, both reviewers pointed out the artificial band resulting from the sounding, please include an explanation in the text, noting why this leads to such an artificial anomaly.

From the additions in the text, it is still not clear, why the control run is objectively the best run, especially with regard to the sensitivity tests in experiments B and C. While the control run itself is compared to measurements – albeit not independent ones – the experiments A, B and C are only compared to the control run. This does not allow us to see, whether this difference actually results in an improved or worsened performance with respect to the measurements. Hence it also does not provide the information necessary to determine that the control run performs better.
If the goal is to show that the control run is optimal, then comparing all experiments to the measurements would make more sense. Since this modified version of WISSDOM includes more input data sources (i.e. more constraints in the cost function), it is valuable to reassess, whether the choice of weights (experiments C) still performs well. The parameter choice for the integration of the AWS data (experiments B) seems to be highly dependent on AWS location and topography and hence also requires a verification against measurements here.
In conclusion, I would suggest to show the results of the sensitivity tests in experiments A, B and C with respect to the sounding and the QVP and adapt the discussion accordingly.
With respect to the verification measurements being used in the algorithm, it is important to state clearly in the text that these are not independent measurements and that the control run is not verified independently. The further experiments A can then show, how WISSDOM performs against soundings, when soundings are not used (as an example), or against QVP, if LIDAR is not used.

Specific remarks:

The following line references highlight minor issues. The line numbers refer to the revised, marked-up manuscript.

Line 32: automatic weather stations (AWS) – throughout the manuscript it is often written AWS station, which is redundant. Please remove “station” after the acronym throughout.

Line 62: It would be interesting to add the characteristic scales to the examples of weahter systems.

Line 78: Based on… - This is a grammatically incomplete sentence.

Line 96: Performing immersed… - Please revise the language.

Line 122ff: This sentence does not make sense grammatically.

Line 138: I would suggest to rephrase to: *A resolution of 50 m was chosen in this study, as the Doppler lidars’ respective horizontal resolution averages 40-60m.* The sentence discussing the Doppler radar seems a bit out of place here, what would you like to highlight with this?

Line 159ff: Please revise the language.

Line 168: Please include the description of the gradient descent technique to converge towards a solution. Also, it is not a random guess, if it is based on the available background data, or set to 0, please clarify this in the text.

Line 211: How does WISSDOM perform thermodynamic retrievals?

Line 224: As the sounding and reanalysis data are not available on a 12 min resolution, how is the time constraint performed? You describe that soundings are launched 3-hourly and the reanalysis data is also provided at a 3h timescale.

Line 289: RHI is not defined.

Line 377: My previous question rather addressed the choice of horizontal and vertical resolution, than that of choosing a Cartesian coordinate system. How did you determine your grid parameters? Did you center your gridboxes to the AWS where possible? Is the horizontal and vertical resolution primarily determined by the LIDAR data?

Line 425: I disagree with this statement. The analytical solution for WISSDOM neither prescribes the spatial integration of the AWS nor the weights, from my understanding.

Line 433ff: These weights were determined for Doppler data at a different spatial resolution and for less additional observational constraints. Considering this, the weights should not be seen as optimal per se and reevaluated in experiments C, with respect to the measurements. These sentences also have some grammatical errors.

Line 471: This description of the QVP is still insufficient to understand fully, how it is derived.

Line 509: Please revise the grammar.

Line 535: The lower correlation here may also stem from the QVP method, please elaborate. “Less coverage” instead of “less coverages”

Line 597f: Please revise the phrasing.

Line 605: Please revise the phrasing.

Line 629ff: I do not understand the conclusion here, please rephrase.

Line 663: This sentence is grammatically incomplete.

Line 668: Revise to “the circular artefact is removed when increasing VE to 90 %.”

Line 732: The conclusion here is unclear. If the weights are insensitive, does this suggest that the data source does not have a large impact on the outcome?

Fig 16: Please mention the artefact from the sounding explicitly in the text.

Fig 18 b): W for the lidar QVP is not mentioned in the caption.

Line 791ff: Please revise the grammar.

Line 806f: Please revise the grammar.