JGR-2022-218
Responses (highlighted in red) to Referee#1
16 Dec. 2022

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.

We appreciate Referee#1 providing helpful and insightful comments in this round, which help us to improve the manuscript substantially. We have carefully checked your comments, and more analyses and discussions associated with the performances of the WISSDOM synthesis have been added to this revision. In addition, we put the statements regarding the non-independent evaluations on the results of the control in this manuscript. A set of responses to your comments is provided below. Specific locations of modified portions (marked as underlines) were also noted as the number of lines in the revised manuscript.

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.

We have checked the problematic grammar carefully throughout the manuscript. The native speaker has also revised the texts throughout the manuscript.

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.

1. The main difference between J2 and J6 is that the uniform wind speed and wind direction (only one value of u-wind and v-wind) was adopted at each level in the entire WISSDOM domain in J2. However, the modified WISSDOM used various values of wind speeds and wind directions along the sounding tracks; any number of sounding observations can be easily merged in the synthesis domain simultaneously in J6. Therefore, we have emphasized these descriptions in the revised manuscript:

   L218-223. The main difference between J_6 and J_2 is that the sounding data with various wind speeds and directions were used as an observation for given 3D locations in J_6 instead of the constraint of homogeneous background winds (i.e., uniform wind speed and direction) for each level in the studied domain in J_2. An additional benefit of J_6 is that any number of sounding observations can be efficiently adopted in the WISSDOM synthesis domain

2. Different spatiotemporal resolutions of the inputs may cause problems applying the same interpretation method. More efficient ways and flexible interpolation methods are desired for the individual inputs. One of the solutions is to put the inputs into individual constraints, and then the constraints can be further applied with suitable interpretation methods. In addition, individual constraints (ex., the AWS) considered the time if the temporal resolution of inputs is equal to or higher than the time intervals of the WISSDOM outputs. We have expanded the explanations about this issue in the test as follows:

   L174-182. However, various datasets with different spatiotemporal resolutions are not favorable for appropriate interpolation of given grid points of WISSDOM synthesis, and the accuracy and reliability of the background may have been significantly affected by such a variety of datasets. Thus, these different observed or model data should be treated differently to minimize uncertainties and improve accuracy. Therefore, one of the improvements in the modified WISSDOM is that these inputs were individually separated into independent constraints with flexible interpolation methods. In addition, individual constraints considered the time if the temporal resolution of the inputs was equal to or higher than the time interval of the WISSDOM outputs.

3. We have mentioned that the WISSDOM is a kind of gradient descent technique to converge toward a solution in the text in:

   L135-137. The 3D winds were derived by variationally adjusted solutions to satisfy the constraints in the cost function; thus, this is a gradient decent technique to converge toward a solution.

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.

We now add a paragraph that describes the lidar QVP wind retrieval. We used the VAD technique. Assuming no horizontal divergence, the vertical wind can be estimated with a PPI scan with a high elevation angle. This study uses the lidar QVP retrieved from the lidar at the DGW site only.

L452-461. The QVP of horizontal and vertical winds were retrieved based on the so-called velocity-azimuth display (VAD) technique (Browning and Wexler, 1968, Gao et al., 2004). We regressed the Fourier coefficients of the Doppler velocities of the 80° PPI under the linear horizontal wind assumption and obtained the horizontal wind profile. The vertical wind was retrieved under the assumptions of constant vertical wind, zero terminal velocity of aerosol particles, and no horizontal divergence [see Kim et al. (2022) for details on the wind retrieval]. The accuracy of the retrieved wind profile is suitable for the WISSDOM wind evaluation, given the low root mean square deviation (RMSD) of < 2.5 m s^{-1} and high correlation coefficient of > 0.94 of horizontal wind speed as shown in the comparison against 487 rawinsondes (Kim et al., 2022).

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.

It may have been produced by the large gradient between each input when their weights were set to be the same, and the results of the WISSDOM synthesis cannot converge well. The possible explanations for this artificial band have been provided in the text.

L722-728: Notably, significant variances usually existed between the observations and reanalysis datasets due to various spatiotemporal resolutions. The results of scenario C-3 do not converge well because there was a relatively more significant gradient between each input as their weighting coefficients were set to be the same (i.e., 10^6). In this way, the effects of poor convergences might be amplified and superposed with the AWS and lidar observations along the sounding tracks. This may be a possible reason that artificial signals existed over the DGW site in scenario C-3.

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.

Thank you very much for this valid point and valuable suggestions. Although the results of the control run provided plentiful information on the discrepancies in the studied domain for each experiment, it takes work to evaluate the difference in facts. Therefore, we have added more analysis relying on your suggestions. The discrepancies between sounding observations along its tracks and QVP above the DGW site and the results of each experiment have been well documented and discussed in this revision. Also, we have declared a statement as the control run is not verified independently with the observations at the beginning of Section 4.2, where we started the analysis in the intercomparison between the derived winds and observations. We also provided some suggestions on the setting in the WISSDOM synthesis. Please refer to the modifications:

L465-468. Because the verification observations are being used in the WISSDOM synthesis, the results of the control run are not verified independently; nevertheless, detailed discussions regarding the results of the sensitivity tests for the observations are presented in Section 5.

L636-650. In addition, the discrepancies in derived 3D winds between sounding observations and QVP were also examined along the sounding tracks (Fig. 13b) and above the DGW site (Fig. 13c). Sounding observations played an essential role in the derived winds along its tracks. The maximum discrepancies of u- (v-) component winds are exceeded by approximately –2 (–1) m s^{-1} if the WISSDOM synthesis lacks sounding observations. However, small discrepancies (nearly 0 m s^{-1}) were presented when the sounding (lidar) data were (not) implemented at all levels in A-1. The peaks in the discrepancies manifested the potential impacts from the lidar and AWS. This may be a result of lidar (AWS) having relativity higher data coverage at ~1.4 (0.8) km MSL (cf. Fig. 4). The maximum discrepancies between the derived winds and the QVP winds are approximately –4 and 4 (–1 and 0) m s^{-1} associated with u-, v- (w-) component winds. Generally, the results reveal similar trends in A-1~A-5, which also implies that all the inputs in the WISSDOM synthesis are equally significant against the QVP. In summary, the results of this experiment (cf. Fig. 13) show that the lidar, sounding, and AWS data are more critical inputs in modified WISSDOM. Therefore, it will be beneficial if various inputs can be included in the synthesis.

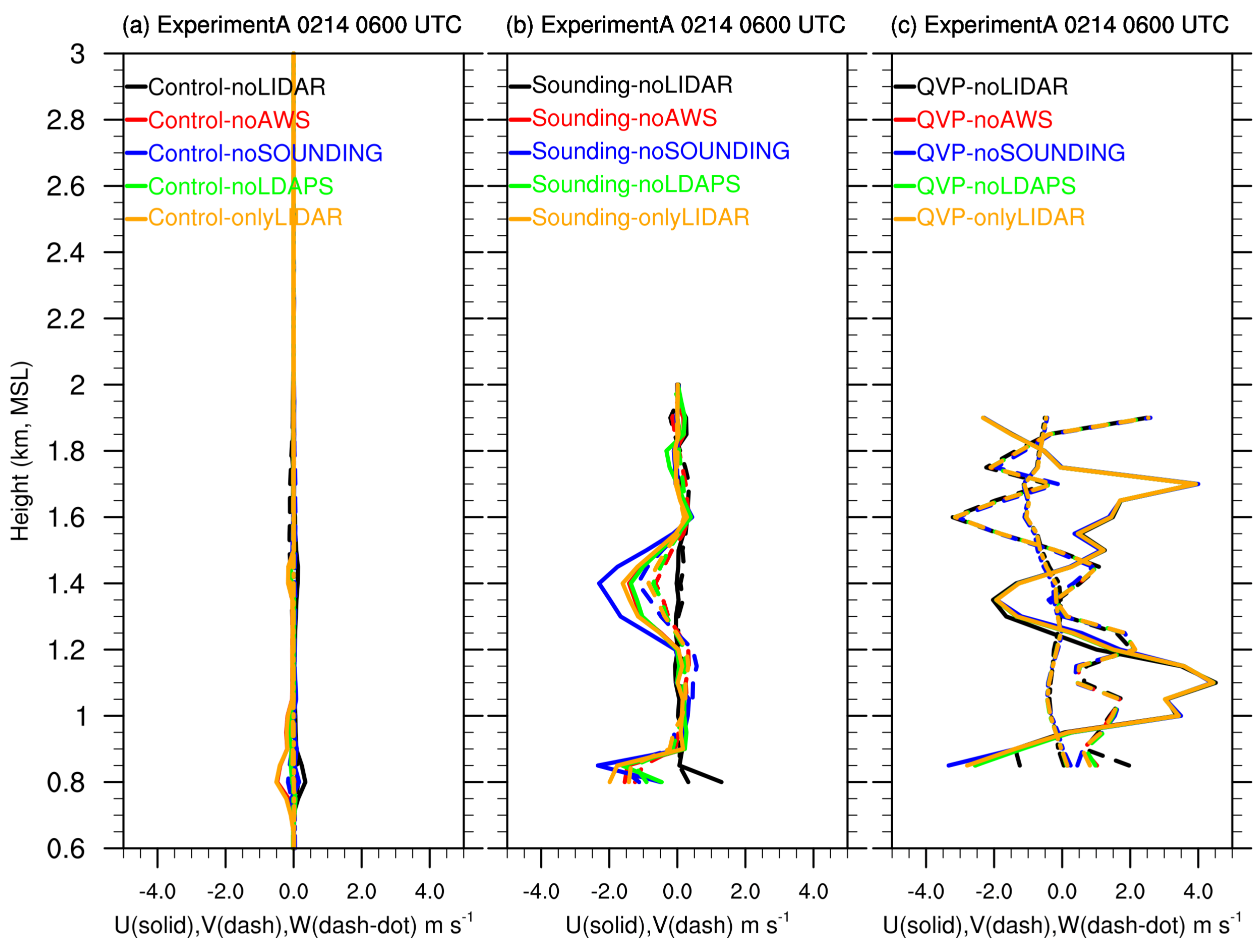

Figure 13. (a) Vertical profiles of averaged discrepancies of 3D winds for each design in Experiment A at 06:00 UTC on 14 Feb. 2018. The averaged discrepancies of u-, v- and w-component winds were plotted by solid, dash, and dash-dot lines, and the black, red, blue, green and orange lines indicate A-1, A-2, A-3, A-4 and A-5, respectively. (b) The same as (a) but for the discrepancies of sounding observations and u-, and v-component winds. (c) The same as (a) but for the discrepancies of QVP.

L696-705. Figs. 15b and 15c show the discrepancies of derived 3D winds between the sounding observations and QVP. Their patterns are similar to A-1~A5 (cf. Figs. 13b and 13c), except there were relatively larger values of u- (v-) component winds at lower layers (approximately –3 and 1 m s^{-1}) in B-1 (Fig. 15b). The v-component winds also presented larger values (exceeded ~3 m s^{-1}) below ~1.2 km MSL compared with the QVP (Fig. 15c). The conclusions indicated that the moderate setting (i.e., RI is 1 km) would be helpful to obtain minor differences with the control run, sounding observations and the QVP. On the other hand, the limited setting in experiment B (i.e., B-1) was helpless. In addition, the wind directions and speed should be dominated by terrain, and the implementation of AWS data is crucial for the modified WISSDOM synthesis, especially in the lower layers.

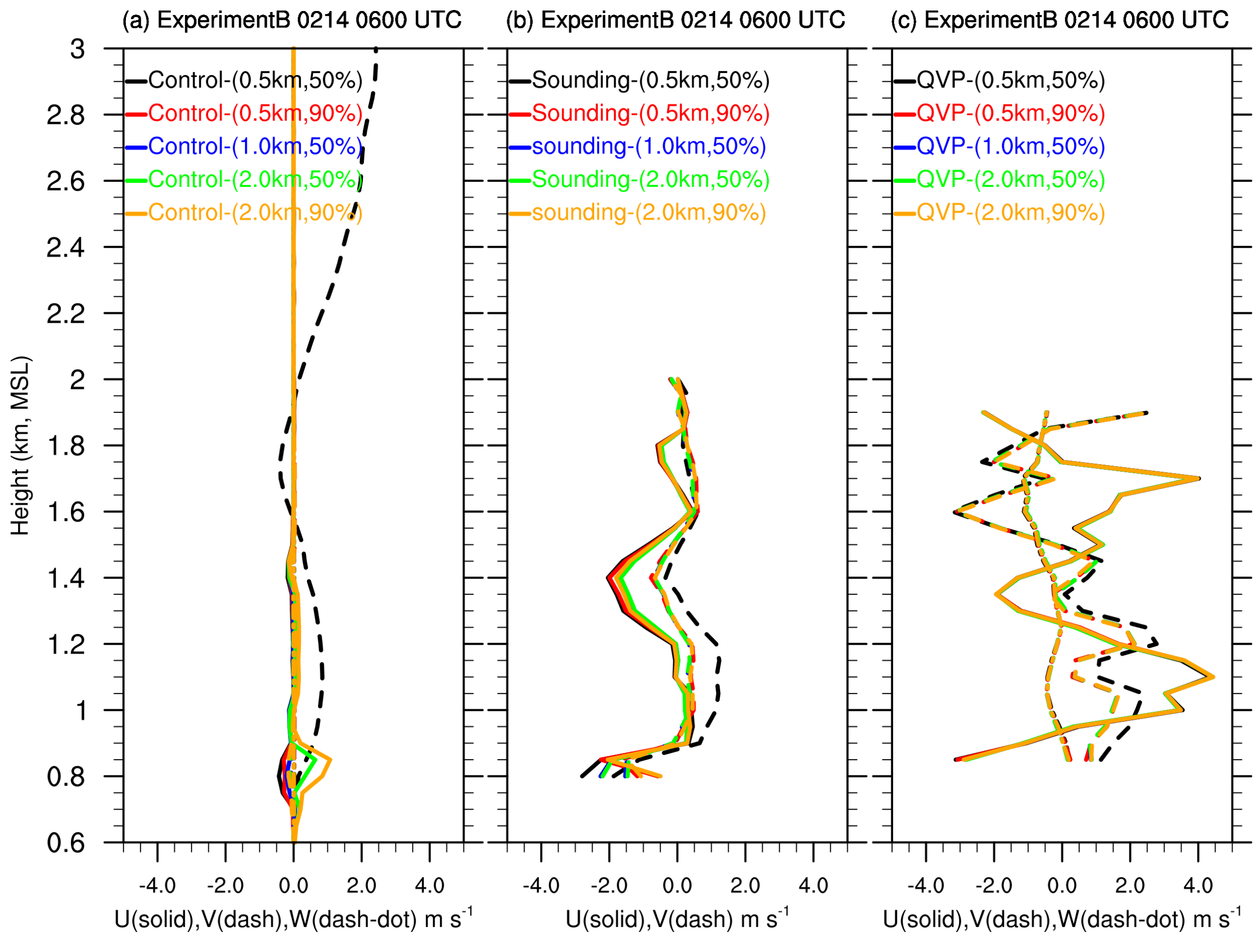

Figure 15. The same as Fig. 13. but for B-1~B-5.

L740-748. The discrepancies in derived 3D winds between sounding observations and QVP in Experiment C were also examined. Compared to the sounding observations, more significant discrepancies in the u- and v-component winds (exceeded ~20 m s^{-1}) can be obtained when reducing (increasing) the weighting coefficients of the AWS (LDAPS) data (Fig. 17b). However, the impacts of lidar against the QVP are shown; their discrepancies are in the range of –1 to 2 m s^{-1} for the u-component winds in C-2 (Fig. 17c). The conclusions reveal that the weighting coefficients of the AWS and LDAPS (lidar) are (moderately) significantly sensitive to the derived winds. Therefore, the weighting coefficients of LDAPS and AWS are not necessarily changed much.

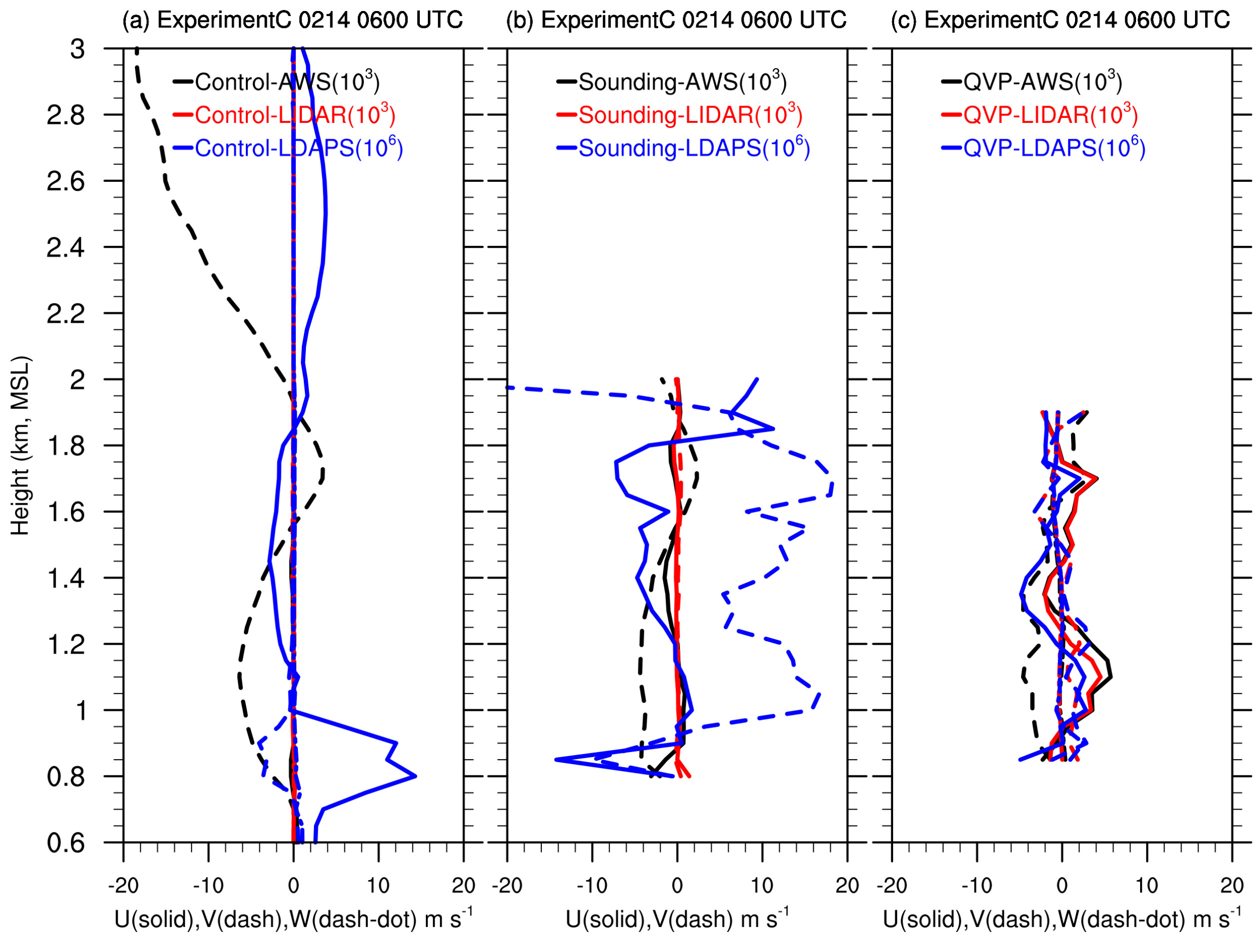

Figure 17. The same as Fig. 13 but for C-1~C-3.

Specific remarks:

The following line references highlight minor issues. The line numbers refer to the revised, markedup 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.

The redundant words have been removed throughout the manuscript.

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

Thank you for the suggestion. As mentioned in the text, we have added characteristic scales to every weather system and phenomenon. The descriptions of this can be seen in the following:
L47-53. Most comprehensive applications of the derived winds were adopted to document kinematic and precipitation structures associated with various weather systems or phenomena at different scales from thousands, hundreds, and a couple of kilometers, such as cold fronts, typhoons, tropical cyclone rainbands, convective lines, and nonprecipitation low-pressure systems (LPS) (Yu and Bond, 2002, Yu and Jou, 2005, Yu and Tsai, 2013, Yu and Tsai, 2017, Tsai et al. 2018, Yu et al., 2020, Cha and Bell, 2021, Tsai et al., 2022).

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

This sentence has been corrected as:
L66-67. Most developed techniques are based on the contexts of weaknesses from the above schemes on wind retrievals.

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

The language has been revised.
L80-83. Liou and Chang (2009) is the first purposes of this algorithm. Furthermore, they performed immersed boundary method (IBM, Tseng and Ferziger, 2003) in WISSDOM, and its scientific applications were documented in Liou et al. (2012) and Liou et al. (2016), respectively.

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

This sentence has been corrected as:

L106-112. However, the original WISSDOM only provided 3D winds under precipitation conditions. It does not work well under clear-air conditions because Doppler radar cannot easily detect radial velocity without precipitation particles. To obtain high-quality 3D winds under clear-air conditions, the radial velocity observed from the scanning Doppler lidars can be used in WISSDOM. The results will allow us to investigate the initiations of precipitation systems in advance of rainfall and snowfall, which is an essential benefit rather than Doppler radar in related research topics.

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?

Revised as suggested, and the redundant parts have been removed.
L122-123. A resolution of 50 m was chosen in this study, as the Doppler lidars' respective horizontal resolution averages 40-60 m.

Line 159ff: Please revise the language.

The descriptions have been revised as follows:
L141-146. Liou et al. (2012) applied the IBM in WISSDOM to consider the effect on the nonflat surfaces. One of the advantages of IBM is providing realistic topographic forcing without changing the Cartesian coordinate system into a terrain-following coordinate system. More scientific documentation associated with the interactions between terrain, precipitation, and winds in different areas can be found in Liou et al. (2016) for Taiwan and in Tsai et al. (2018) for South Korea.

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.

The descriptions have been added and clarified in the text:
L135-137. The 3D winds were derived by variationally adjusted solutions to satisfy the constraints in the cost function, thus this is a gradient decent technique to converge toward a solution.
L150-151. Note that the $\mathbf{V}_t$ will be first guessed, resulting from the background of the sounding observations used in this study.

Line 211: How does WISSDOM perform thermodynamic retrievals?

The descriptions have been revised clearly:
L193-196. The main advantage is that using vertical vorticity can provide further improvement in winds and thermodynamic retrievals from a method named as Terrain-Permitting Thermodynamic Retrieval Scheme (TPTRS, Liou et al. 2019).

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.

We used the same time constraint if the temporal resolution of inputs is lower than the synthesis time step. The descriptions have been added to the text as follows:
L210-213. Notably, relatively minor changes in environmental conditions were assumed in WISSDOM due to the limitation on the coarse temporal resolution from specific inputs. For example, the closest time step of a sounding observation or LDAPS dataset was chosen regarding the synthesis time, and the time constrain was set to be the same.

Line 289: RHI is not defined.

The RHI has been defined in the text (L98).

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?

Yes, the LIDAR data determined the horizontal and vertical resolutions. It is because we are trying to get derived winds in a higher resolution, and the lidar can provide the information with the highest spatiotemporal resolution of observations. We determined the given grid points based on the characteristics of lidar data in this study. Therefore, we did not center our grid boxes to the AWS where possible, but we have interpolated all data resources on the given grid points. There were several descriptions related to this issue, and we further emphasized that in the text:
L122-123. A resolution of 50 m was chosen in this study, as the Doppler lidars' respective horizontal resolution averages 40-60 m.
L359-360. In this study, the horizontal and vertical resolutions of given grid points were primarily determined by the characteristics of 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.

This is a valid point since our goal is only to show that the control run is optimal. Therefore, the redundant descriptions have been removed from the text.

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.

Since we had no information about determining the weights before we applied the lidar data in the WISSDOM synthesis, the setting of weight in WISSDOM by using radar is the only reference. Therefore, we first decided on the weights and followed

the concepts from the previous studies as more (less) weights in observational (reanalysis datasets and modeling) inputs. Consequently, we are trying to find the optimal setting for the weights. So, it is a reason why we need to do sensitivity tests to understand the possible variances associated with different weights. We have emphasized this statement in the following:
L417-419. Consequently, sensitivity tests were performed to better understand the possible variances associated with different weighting coefficients when the lidar data were implemented.

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

Please refer to our responses to your general remarks 3.

Line 509: Please revise the grammar.

The grammar of this sentence has been corrected as:
L502-505. Only relatively larger IQR (between ~–5 and 5 degrees) and larger median values (between ~0 and 5) can be found at the lowest level. The interquartile range (IQR) and median values of the wind direction differences are smaller (between ~0 and 2.5 degrees) during the entire research period (Fig. 8a).

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

The word has been corrected.

Line 597f: Please revise the phrasing.

The sentences have been revised as follows:

L582-584. In this section, the impacts of various datasets on data implemented in the WISSDOM synthesis were evaluated. In particular, the quantitative variances between each design, control run, sounding observations, and the QVP can be estimated.

Line 605: Please revise the phrasing.

The sentences have been revised as follows:
L592-594. An additional test was designed as only Doppler lidar data are used without other constraints from $J_6 \sim J_8$ (A-5) to evaluate the performances between the modified and original versions of WISSDOM.

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

The sentences have been revised as follows:
L615-618. Relatively weak winds were presented near the surface from the results of A-5 (Figs. 12i and 12j). These results reflect that the additional constraints play crucial roles, especially at lower layers. Furthermore, it is implied that the winds can be reasonably retrieved when additional constraints are set in the modified version of WISSDOM.

Line 663: This sentence is grammatically incomplete.
The grammar of this sentence has been corrected as:
L659-661. Because the average distance is approximately 0.1 to 2 km between each AWS site, there were five designs (B-1~B-5) in Experiment B with ranges of RI (VE) between 0.5 km (50%) and 2 km (90%).

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

The sentences have been revised.

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?

This paragraph has been rewritten based on the new analysis:
L740-748. The discrepancies in derived 3D winds between sounding observations and QVP in Experiment C were also examined. Compared to the sounding observations, more significant discrepancies in the u- and v-component winds (exceeded ~20 m s^{-1}) can be obtained when reducing (increasing) the weighting coefficients of the AWS (LDAPS) data (Fig. 17b). However, the impacts of lidar against the QVP are shown; their discrepancies are in the range of –1 to 2 m s^{-1} for the u-component winds in C-2 in (Fig. 17c). The conclusions reveal that the weights of the AWS and LDAPS (lidar) are (moderately) significantly sensitive to the derived winds. Therefore, the weighting coefficients of LDAPS and AWS are not necessarily changed much.

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

Please refer to our responses to your general remarks 4.

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

The caption has been revised.

Line 791ff: Please revise the grammar.

The grammar of this sentence has been corrected as:
L804-809. Relatively reasonable winds can be derived with optimal settings in modified WISSDOM, and the setting of 90% (50%) in VE is also recommended over complex terrain (flat surface). These sensitivity tests will help verify the impacts against various scenarios and observational references in this area. The conclusions

can also be a good reference for deciding the best locations to deploy the instruments.

Line 806f: Please revise the grammar.

The grammar of this sentence has been corrected as:
L819-823. The detailed wind structures can be well documented for any meteorological phenomena in clear-air conditions (e.g., land–sea breezes, microdownbursts, nonprecipitation low-pressure systems, etc.) via a modified version of WISSDOM. It also has broad applications in site surveys of wind turbines, wind energy, monitoring wildfires, outdoor sports in mountain ranges, and aviation security