red: point-by-point response
blue: relevant changes made in the manuscript

Replies to the editor's comments

1) Referee comment: make the quality-metrics response fully explicit
- Please revise your response letter to explicitly list the indicators now described in the manuscript and state that they are available at the level of individual time–height bins (and/or individual profiles).

We revised the response letter. You can find the revised version of the replies to referee comments below.

- Please also add 2–3 sentences in the manuscript explaining how these diagnostics are intended to be used in data assimilation (e.g., QC flags and/or weighting/error inflation guidance).

The retrieved wind vectors in the final dataset are accepted or discarded based on the provided QC flags. The standard module chain provides QC flag thresholds (listed in Tab. A1), which discard unreliable wind vectors, as demonstrated by the radiosonde comparison. The thresholds can be modified by the users if desired. Since quality controlled wind profiles are provided by the software, no additional QC is needed for data assimilation. Typically, an instrument type specific standard error is considered for data assimilation (typically on the order of 1 m/s for Doppler wind lidar). The deviations observed in the radiosonde comparison confirm the plausibility of the assumed standard error for Doppler wind lidars. If desired, data assimilation could consider the QC indicators to perform weighting of the retrieved wind vectors. To our knowledge, such a weighting of observations has not been performed for Doppler lidar data up to date. Possible methods include weighting of the retrieved wind vector by the residual velocity variance or spectral width, which are indirect measures of turbulence within the retrieval volume.

We added statements detailing the potential use of the QC indicators in data assimilation to the manuscript.

2) Code and Data availability: Please provide a permanent link for measurement data from Neumayer Station and Payerne.

We requested a DOI, but unfortunately DOIs are currently not available for raw measurements. To make these measurements permanently available, we uploaded the relevant Doppler lidar measurements from Neumayer Station and Payerne to the dataset associated with this study `https://doi.org/10.5281/zenodo.14844887`.

We added the data to the permanently available dataset associated with this study `https://doi.org/10.5281/zenodo.14844887`.

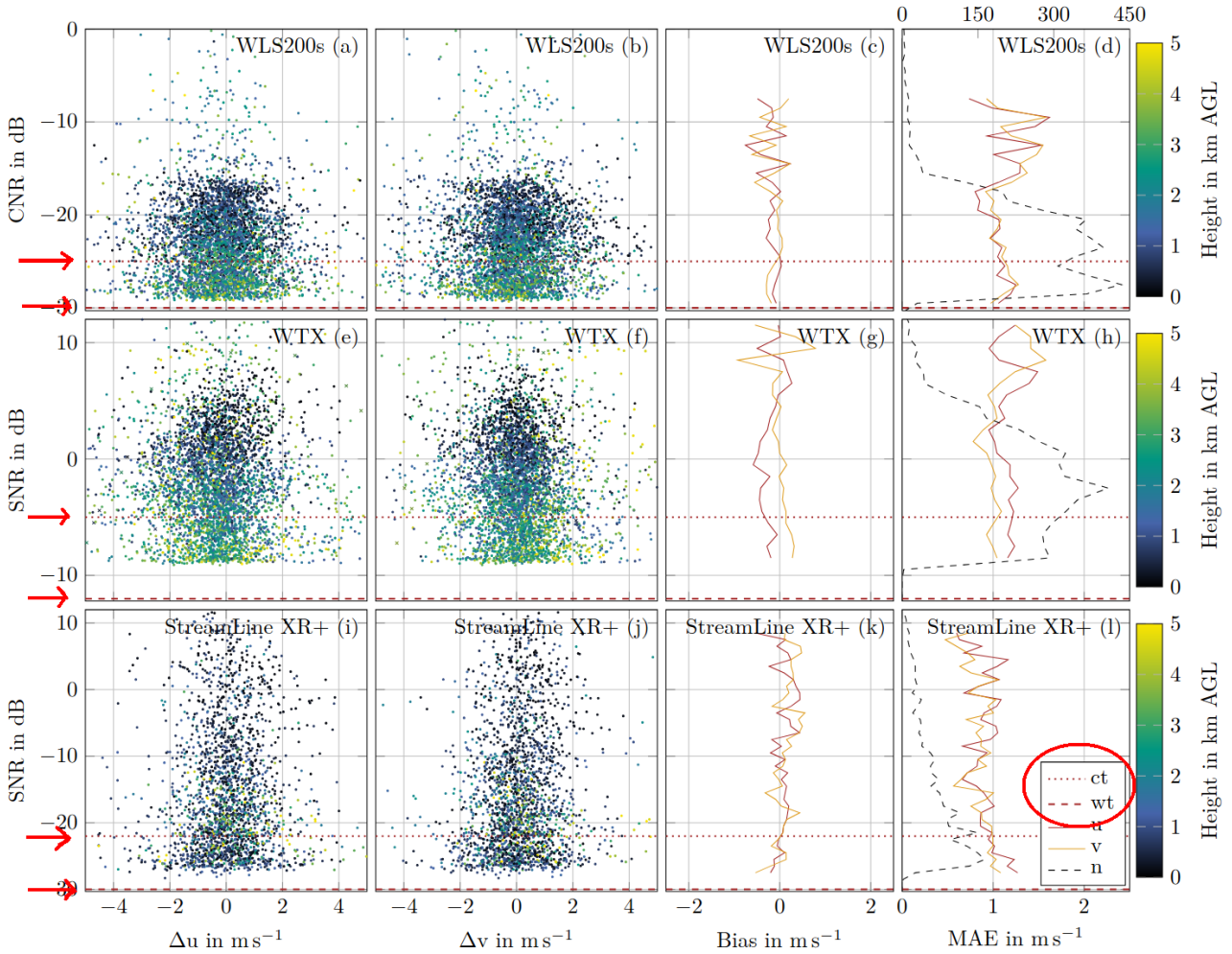
3) Figures: color-vision deficiency (CVD) accessibility Please ensure all figures are interpretable for readers with CVD and in greyscale reproduction. You can check the GMD requirement at
https://www.geoscientific-model-development.net/submission.html#figurestables.

We checked all plots and assume you are pointing on an ambiguity of the colours used for u and v. Hence, we revised the colourbar and replaced all concerned plots.

We revised the colourbar and updated the concerned Figures.

4) One more comments from a reviewer: In figure 16 the 'conservative thresholds' (ct) for CNR/SNR could be marked, and it should be noted that the 'weak threshold' (wt) is in all subplots the lowest value of the ordinate. The low CNR acceptance based on the used image restoration method occurs only between these two thresholds. The errors in the range between ct and wt should be discussed: Is there a systemtic increase in bias and MAE in this region.

We replied to this comment in our previous response to the reviewer, see below. We added the suggested thresholds to Figure 16 as suggested and added the suggested extension of the discussion in the text. Potentially there was confusion with Figure E4 which looks very similar, but shows the distance on the ordinate instead of the CNR and includes, thus, no CNR thresholds.



Addressed in the previous reply and changed in the revised submission.

Replies to referee comments

1. In figure 16 the 'conservative thresholds' (ct) for CNR/SNR could be marked, and it should be noted that the 'weak threshold' (wt) is in all subplots the lowest value of the ordinate. The low CNR acceptance based on the used image restoration method occurs only between these two thresholds. The errors in the range between ct and wt should be discussed: Is there a systemtic increase in bias and MAE in this region.

We thank for this suggestion. We slightly extended the ordinate and marked both thresholds in Figure 16. We added the proposed discussion in the previous revised version.

Added the thresholds in the Figure and the discussion in Section 5.3 in the previous revised version.

2. I understand that properly quantifying uncertainty in Doppler lidar measurements is very difficult as it depends upon not only the measurement of LOSV, but sampling, and assumptions of homogeneity. Having said that the purpose of the paper is to have wind retrieval algorithm for data assimilation which requires information about uncertainty in individual profiles. So, I would think having some metric that informs the quality of individual profile or individual height bin would be very helpful.

We thank the referee for the clarification. In the last reply we had a focus on uncertainty as a quantity, not as a quality indicator and apologize for the misunderstanding. We have indeed some suitable indicators and agree that they become important for quality metrics.
In principle, the retrieved wind vectors in the final dataset are accepted or discarded based on the provided QC flags. The standard module chain provides QC flag thresholds (listed in Tab. A1), which discard unreliable wind vectors, as demonstrated by the radiosonde comparison. The thresholds used for filtering can be modified by the users, if desired.
To provide further insight into the measurement quality, the CNR and the spectral width provided by the

instruments for each single measurement are evaluated and added for each bin as quality indicator to the level 2 dataset. This is carried out twice: (1) considering all measurements contributed to the finally retrieved wind vector, and (2) considering all measurements associated with the bin. To provide insight into flow homogeneity within the retrieval volume (e.g. due to turbulence), the variance of the residuals is added to the level 2 dataset as a quality indicator. To provide insight into the measurement and retrieval availability, indicators as the number and share of considered measurements as well as the CN can be added to the exported level 2 dataset. We added this information at the end of Section 4.1.

Added a paragraph on the quality metrics and indicators at the end of Section 4.1 as suggested by the editor.