We thank the reviewers for the relevant comments that helped to improve the manuscript.

Answers to reviewer #2

General comments

Although the campaign was carried out in July 1999 and a UV lidar based on the double-edge technique concept tested in the campaign is to be soon flown on the ESA's ADM-Aeolus mission, the methodological aspects are still valid, especially in the prospect of calibration-validation activities for that mission. Although this fact is mentioned by the authors (e.g. p. 4554, lines 13-15; p. 4563, lines 1-3) it should be further stressed and probably included in the abstract.

Agreed, as the paper can be useful after ADM/Aeolus launch for ground validation.

The title highlights the 0.355 um lidar, but in the body of the paper this lidar seems to be on an equal footing with the other lidar instruments. The authors should stress the special role this lidar played in the campaign – was the campaign decisive for the selection of a UV double-edge direct detection lidar for ADM-Aeolus? Was the decision already made and the campaign intended only to confirm the soundness of the technological choice? Moreover, in the reviewer's understanding the selection of a UV lidar for ADM-Aeolus is based on its capability for measuring wind velocities in free atmosphere regions where molecular scattering is much stronger than aerosol scattering, which involves scattering greatly enhanced as the wavelength is decreased, as well a wide spectrum of the scattered radiation because of the wide spectrum of velocities of the molecules around their mean (wind) component. This seems not to be stressed enough; in fact the sentence that addresses this, starting on line 25 of page 4553, – "For atmospheric molecules are uniformly distributed geographically with a known dependence in height, ESA decided to select in 1999 a spaceborne wind lidar based on molecular scattering at 0.355 µm" – is probably too weak and a little confusing, as there is probably not other way for measuring the wind velocity in the upper atmosphere through backscatter lidar techniques than relying on molecular backscatter.

The VALID campaign was conducted in the summer of 1999 while the decision to select 2 of the 4 ESA's Earth Explorer Missions will be taken few months later in October 1999. As said in the paper, the UoG's 0.355-µm Wind lidar was still in a testing phase during VALID. So, only preliminary results from the 0.532-µm DDLidar and 10.6-µm HDLidar were presented during the selection contest in Granada in October 1999. At the time, the objective was to show that wind-sensing lidars provide the same answer when operated side-by-side and provided the SNR are sufficient. For lidar, SNR relies on atmospheric backscatter and instrument design (laser energy x receiver area)

Specific comments

Because a UV Doppler lidar should perform better in the absence of aerosol than lidars *at longer wavelengths (at least for similar energy of the transmitted pulses and collecting-aperture size), one would expect a priori that field campaign results reveal better performance above the boundary layer for this lidar than for the other ones, or at least an improvement in its comparison with the 0.532 mm one. However an analysis of the performance as a function of the height is not found in the paper. Tables 4 and 5, which summarize the performances taking as reference the radiosonde-derived velocity, do not state the range of altitude values for which Eq. (2) and (3) have been computed. The range must be stated in the table caption as well as in the text, where it doesn't seem to be specified either. It is also strongly suggested that results are given for different range intervals, to show possible changes in the performance as a function of height. Likewise, more detail should be given about the 0.532 mm lidar direct-detection Doppler lidar to place in full context the comparison with the 0.355 mm one. In fact, there is a lack of homogeneity in the way the different lidars are described, much more detailed technical information being given in the text for the 10.6 mm heterodyne one than for the 0.532 mm one; even if specific references are given for the latter, more technical details should be explicitly included in the paper. In turn, the characteristics of the 0.355 mm lidar are given in table form (Table 1).*

We agree that dividing the analysis into lower and upper troposphere will help greatly the general understanding. The 0.532-µm lidar description and technical specifications will be inserted.

In general, a critical assessment of the campaign results is missed, as well as the reasons for the selection of those that are highlighted: how well are the objectives stated on lines 24-27 of page 4554 (to compare the performances of the different lidar techniques in various meteorological conditions, to demonstrate that the retrieved wind velocities are the same (within the statistical error) and to explain the differences in complex situations) fulfilled? Some more discussion about this would be helpful for the reader. What's the reason for selecting the 12 datasets of table 3? Their representativeness of different atmospheric conditions? Are the results from other possible datasets in agreement with those of one of those selected corresponding to similar atmospheric conditions? What's the reason for choosing to highlight the results of 20, 21, and 22 July rather than those of other selected days? In connection with Fig. 4, would the other lidars in the campaign, in particular the 0.532 mm one, show similar deviations beyond the instrument accuracy limits? Concerning these limits, Table 5 gives two values for the instrumental error for DC-DDL and DEDG: what do those two values correspond to? There seems also to be some internal inconsistence in the ascription of the cause for those deviations: on page 4561, lines 17-20, the reason seems to be attributed to a change in the aerosol properties, whereas on page 4562, lines 19-20, the explanation seems to be the difference between the volumes sensed by the lidar and the ones where the radiosonde has been drawn by the wind.

Even if the paper is focused on the performances of a 0.355 µm direct detection Doppler wind lidar developed at University of Geneva and not on the results of Valid campaign, we agree to add some more discussion to the results of the campaign. The 12 selected datasets are firstly selected when the 0.355-µm DDlidar was operational, and secondly representative of the variability of the meteorological conditions (clear-sky, overcast, strong wind, foggy, high aerosol load…). As said before, DEDG instrument suffered from misalignment during transport and the datasets were also selected relatively to data availability.

The selected days (20, 21 and 22 of July) represent cases, where it is noticed a discrepancy between the instruments and the radiosounding. These particular cases may explain the errors and their causes. For example, in Figure 3 it is shown how the different spatial resolution plays a role in arising discrepancies among the instruments, as some wind fluctuations are smoothed. 22nd July put in evidence how aerosols and lower clouds contaminate DEDG measurements, even if the instrument was tuned "ad-hoc" to desensitize the measurement from aerosols. A similar DDL plot as Fig. 4 would be interesting but beyond the scope of the present manuscript, that is focused on DEDG performances and development. Figure 4 shows that the DEDG set-up does not work in situations, but depends on aerosol loading and cloud formation. For this reason, in order to desensitize the measurement from aerosol and clouds, a shot-by-shot calibration should be performed.

We agreed that an explanation is missing on how the uncertainty is retrieved on Table 5. It will be integrated in the revised manuscript. The interval is showing the maximum error, depending on measurement range. For DEDG, at 10km, the error is more likely to be 3m/s (calculated over 20s or 600 shots).

From the analysis, it seems that there is not a single cause prevailing among the others for discrepancies between instruments and radiosounding. We agree that the actual version it is not clear and these paragraphs should be rephrased.

The explanations on lines 15-20 of page 4553 about heterodyne detection and direct detection suffer from a too hasty writing. The sentence "Heterodyne detection technique analyzes the backscattered spectrum from aerosol or cloud particles" is imprecise. It would be more exact to say that this technique is better suited for measuring the Doppler shift of the radiation (or the spectrum) scattered by aerosols or cloud droplets, because this spectrum is narrow and doesn't pose too demanding requirements on the intermediate frequency stages of the receiver electronics (that should let pass signals with bandwidths on the order of GHz if the radiation scattered by molecules was to be used). A more detailed (without being exhaustive) description of the direct-detection double-edge technique would also be desirable here, to put forward its advantages with respect to the heterodyne one when it comes to rely on molecular backscatter. This should also help understand entries on Table 1 such as "Number of etalon channels", "Laser etalon separation-locking ch." and "Laser etalon separation atm. Ch." For the overall structure of the explanations it would probably be better to start with a brief description of the spectra backscattered by particulates and by molecules, then explaining the suitability of the different techniques to measure the Doppler shift in situations dominated by either aerosol or molecular backscatter.

The suggested changes are applied; together with a more detailed description of direct detection DEDG technique, to clarify entries of Table 1.

Does the reference list given on lines 2-5 of page 4554 intend to be representative in general? If yes, probably references on CO_2-laser-based lidars developed at NOAA should be included. If not, the context in which the references must be considered should be discussed.

Reference to NOAA lidar is now included.

In addition to produce tables like tables 4 and 5 for different height intervals, it would also be interesting to have tables comparing the different instruments against each other. For example, letting aside the question of the different height intervals, table 4 could be split into three tables, each one giving the correlation coefficient for the corresponding measured velocities between pairs of instruments. In this way, a picture of how each instrument behaves as compared to each other would be obtained, and possible effects related to the radiosonde drifting would be revealed (one would expect better match between measurements on very similar volumes).

Agreed. An analysis into lower (0-5km) and upper (5-15km) atmosphere, is done. Also Table 4 is divided into three tables with correlation coefficient for instrument pairs.

Figures 3, 4, and 5 are very small. This can be a typesetting issue, as they don't get blurred when the view is blown up. However the quality of fig. 2 seems definitely too poor.

The resolution of these figures is improved

Technical comments

We took into account all the suggested corrections if not otherwise specified.

21. Page 4560, lines 14-19: "The wind fluctuations due to orography are then more likely meridional than zonal especially in strong wind conditions (Mistral) as shown in Fig. 3. The effects are expected to be stronger in the lower atmosphere (0–5 km). The instrument spatial resolution is an important variable, especially in the lower troposphere, where atmospheric layers are thin. For these reasons, the remote sensors sometimes do not follow the wind fluctuations."
Comment: the explanations given in this sequence of sentences should be stated in a more precise way. What is the spatial resolution an important variable for? The sentence seems to imply that instruments with better spatial resolution should score better in the figures of tables 4 and 5 – by the way, is the last column (corresponding to the radar) of table 4 right?: the numbers are exactly the same as those in the column to its left. However, letting aside the radar, which is not looking along the same direction as the lidars, the system that seems to perform better (higher correlation coefficient, lower difference with respect to radiosonde) is the 10.6 mm heterodyne lidar, which is not the

one with better spatial or temporal resolution.

The spatial resolution is affecting mostly the cross-correlation coefficient, as it takes into account the difference in shape among the profiles. HDL lidar by the way, is showing a better accuracy as the comparison with this instrument is usually done over lesser points and in the lower troposphere, where there are not larges discrepancies as in the jet-stream for example. This sentence will be implemented in the manuscript.

25. Page 4561, lines 23-24: "Strong winds were inducing gravity waves at low altitudes with significant vertical velocity"
Comment: was this confirmed by the radar, which was able to retrieve the vertical component of the velocity?

Unfortunately no vertical wind velocity measurements by ST-Radar were available. Even if the instrument was operated continuously, the ST radar does not provide wind measurements all the time depending on meteorology.