

**Airborne Lidar Observations of Wind, Water Vapor, and Aerosol Profiles During The  
NASA Aeolus Cal/Val Test Flight Campaign  
Authors: Bedka, Nehrir, Kavaya, and Co-Authors**

**Responses To Reviewer 1 Comments**

**- Title:** The title of the manuscript is “Airborne Lidar Observations of Wind, Water Vapor, and Aerosol Profiles During The NASA Aeolus Cal/Val Test Flight Campaign”. Thus, the reader expects more or less only measurements that are used for Aeolus Cal/Val. However, the comparison of DAWN and Aeolus wind data is only a very small (~5%) part of the manuscript. Neither water vapor profiles nor the aerosol optical properties data is used to compare to Aeolus. Thus, the paper mainly demonstrates what is possible with the payload flown during the campaign (also e.g. regarding boundary layer atmospheric conditions, etc.). Thus, it could be thought to make the Aeolus Cal/val less prominent within the title of the paper manuscript. However, if this was the official name of the airborne campaign it is understood that this has to be kept.

We recognize the validity of your comment, but through discussions with NASA and various presentations and publications already publicly available, the official name of the campaign that you see in the title has already been established. We had hoped to be able to include reprocessed Aeolus wind and aerosol data in this paper, which would have enabled a more comprehensive comparison with DAWN and HALO and greater presence in this paper, but this was not possible given ESA’s reprocessing timeline for the Laser-A data record. Such comparisons will be done in the future after reprocessing. For the reasons here, we choose to keep the title as is.

**- Introduction (general):** What is a little missing in the introduction is, what is new or rather special for the presented research results. Is it the first time that an HSRL and a DWL are flown on the same aircraft? Or is the performance of the used instruments much better than the one shown by other groups? Are there any other airborne activities to CAL/VAL Aeolus for wind and/or rather optical properties? Such kind of information would help to put the presented work in an international context.

In terms of other activities ongoing or recently conducted at NASA regarding airborne wind or optical property profile measurements, with regards to wind, the DAWN instrument is one of NASA’s airborne wind profiling instruments, but the only Doppler Wind Lidar. Other radar instruments such as APR-2 and –3 and EXRAD can retrieve vertical component winds in precipitation regions, but not horizontal winds. It was noted in the text that DAWN was flown in CPEX 2017, and previously in Polar Winds I and II. Flight opportunities are relatively infrequent at NASA, so Aeolus Cal/Val was the next flight opportunity after CPEX 2017. Aerosol profiling instrument flights focused on atmospheric composition research are quite frequent at NASA. The HSRL, HSRL-2, and more recently HALO have flown in many recent missions including SEAC4RS, ORACLES, NAAMES, ACT-America, LISTOS, FIREX-AQ, and ACTIVATE.

In terms of the novelty of these flights, this is the first time to our knowledge that a WV DIAL, HSRL, and Doppler Wind Lidar have flown together on a single aircraft. The LASE instrument flew with DAWN during the 2010 GRIP and collected relative backscatter profiles from aerosol/cloud as well as WV profiles. But this mission was one of the first for DAWN, and it did not perform as well as it did in subsequent missions, especially CPEX 2017 and Aeolus Cal/Val.

We choose not to speculate regarding how DAWN and HALO may perform better/worse relative to other international instruments of their kind.

To address your comments regarding the novelty of these flights, we have included the following sentence in the Introduction:

“To our knowledge, this is the first time that quantitative profiles of aerosol optical properties from a High Spectral Resolution Lidar (HSRL), water vapor profiles from a Differential Absorption Lidar (DIAL), and wind profiles from a Doppler wind lidar were simultaneously observed from a single aircraft”

- **Line 44:** A few acronyms, e.g. ALADIN were already introduced in the abstract. Thus, there is no need to introduce the again. The acronym ESA for the European Space Agency is not introduced, however, used in the citations. Thus, it is recommended to introduce ESA.

Done

- **Line 48-49:** “Aeolus observes...LOS”. This was already mentioned in the sentence before (at least for the wind field) → skip or harmonize these two sentences.

Done

- **Line 66-68:** “lack hor. And vert. resolution” → It could be helpful to mention which resolutions, coverage, etc. would be required for observations to be useful for NWP (including reference).

The thrust of the text you comment on was directed to how current space-borne sensors do not offer the resolution nor precision required to address key process-oriented science questions, and how airborne sensors are required to help fill gaps. We have added the following text to provide an example of wind and moisture measurement requirements to address a “Most Important” ESAS (2017) PBL science question:

“For example, the ESAS (2017) Consolidated Science and Applications Traceability Matrix identifies geophysical observables and their associated accuracy required to address a number of key science questions. A ESAS (2017) “Most Important” question *“What planetary boundary layer (PBL) processes are integral to the air-surface (land, ocean and sea ice) exchanges of energy, momentum, and mass, and how do these impact weather forecasts and air quality simulations?”* requires measurement of 3-D wind vector and moisture profiles every 20 km and 3 hours, with 0.2 km vertical resolution and accuracy of 1 m/s and 0.3 g/kg, respectively, that is currently not attainable from space-borne sensors.”

- **Line 84-85:** Why do you use "" here? In which of the following references is this sentence written?

The quotes were not necessary as this exact quote was not extracted verbatim from any one of these references, so they are now removed.

- **Line 115-118:** Does this mean that the beam shaping optics improve the SNR similarly than having 250 mJ pulse energy instead of 100 mJ, or just that the missing 150 mJ were partly compensated by the beam shaping optics?

We have revised the text with the following that should address your question:

“It originally generated 250 mJ per pulse using a crystal amplifier for the Genesis and Rapid Intensification Processes (GRIP) and PolarWinds I and II campaigns described below. However, this component failed and was removed. This change caused the beam size and curvature entering the beam expander (BEX) to be sub-optimum, lowering the heterodyne mixing efficiency. The amplifier space was used to locate beam shaping optics which restored the optimum beam to BEX coupling efficiency, thereby increasing signal-to-noise ratio (SNR) to a level greater than if the amplifier had been replaced.”

- **Line 142:** Acronym INS was not introduced.

Done

- **Line 172:** This sentence is not completely clear to me. On the one hand, quantitative numbers are missing. How good is the comparison to in-situ data? On the other hand, DAWN is expected to have no measurements at flight level, right? With a laser FWHM of 180 ns (54 m), the first 100-200 m might be influenced by the outgoing laser pulse which would shift your power spectrum towards 0 m/s. Thus, a comparison to in-situ winds at flight level might not be too meaningful. Could you clarify that?

We are reporting on the analysis described by Greco et al. 2020. The first valid DAWN wind retrieval which as you note occurred at some distance below the aircraft was compared with the flight level wind measured in-situ by the aircraft. Though there can be some wind gradients in this 100-200 m layer you mention, the insitu data is constantly available and synchronized with every DAWN profile, serving as a stable, though imperfect reference to estimate DAWN precision at the upper portion of each profile.

- **Line 177-179:** It would be worth mentioning that the DLR lidar performance was determined also for airborne measurements (not from ground). So “sondes” are actually dropsondes and not radio sondes. Maybe this can be clarified in the manuscript.

We have clarified that the DLR comparisons were also with dropsonde.

- **Line 184:** “Unreliably”. What does unreliable mean here? Is only the 10 Hz data not available all time, or are values corrupt (also in the 1 Hz data)?

We have clarified this with the revised statement “Throughout the 2019 Aeolus campaign, the INS/GPS unit attached to DAWN periodically had problems with signal acquisition that resulted in unpredictable drifts in recorded aircraft position and orientation.” The 1 Hz DC-8 INS/GPS was extremely stable and was used in place of the 10 Hz unit dedicated to DAWN.

- **Line 200:** LASE was not introduced...is it an acronym for the precursor of HALO?

Lidar Atmospheric Sensing Experiment is now defined in the text

- **Line 208:** “exceeded all expectations”: Concerning what? Reliability? Data coverage? Accuracy? It would be good to provide a short explanation here.

The following text was added/amended to help clarify the “exceeded all expectations statement”.

Despite serving as the first set of engineering test flights, HALO exceeded all expectations regarding laser reliability, measurement sensitivity, dynamic range, and accuracy and precision (to the extent validated during this mission). The results from the Aeolus Cal/Val campaign demonstrated the first new airborne WV DIAL capability within NASA in over 25 years and provides a new observational tool to the community for future process and cal/val studies.

- **Line 231:** water vapor → WV

Corrected

- **Line 236-237:** Have you ever verified if the WV data precision is Poisson noise limited or if there are other systematic contributions to the random error?

It is difficult to assess the sources of systematic error of the HALO WV data without having a statistically significant validation data set from which to compare the lidar data against.

Comparisons to the limited in situ data sets collected during this campaign and ground based radiosonde comparisons have not identified sources of systematic error and have also indicated that the measurement precision does scale with Poisson statistics to beat down shot noise and electronic noise. We believe a detailed discussion on this subject is beyond the scope of this paper. A validation paper is in preparation that discusses in more detail the random and systematic error of the HALO WV data products.

- **Line 239:** water vapor → WV

Corrected

- **Line 252:** "good agreement" What does "good" and "excellent" mean here? Can you already provide numbers that you are later discussing in section 4.4?

Given the very sparse sampling a detailed statistical analysis is not presented here. The following sentences were amended to present high level statistics.

"Qualitative comparisons, however, generally showed good agreement in the lower troposphere and into the PBL, where the HALO WV profiles resolve the shape and general magnitude of the WV measured by the sonde. Comparisons with the DLH in-situ open path measurement, conducted during a spiral, showed excellent agreement with an average percent difference, above 4.5 km and below 1 km (PBL), of approximately 5%. Statistics between 1 km-4.5 km are omitted here due to the sparse sampling statistics and large variability within the in-situ sampling volume."

- **Line 262-263:** Can you quantify "favorably" here? What means "show promise" in detail?

The following text was amended/added to provide more context to the near surface comparisons.

"Preliminary results using the ocean surface echo compare favorably with the DLH in-situ observations, with an absolute difference of less than 10% from the surface up to 315 m. However, it should be noted that the lowest extent of the aircraft spiral limited the DLH observations to ~200 m above sea level so that the lidar comparisons are also limited to 200-315 m above sea level. As with the full profile comparisons to the sondes from above, the surface echo retrievals generally showed good agreement with the shape and magnitude of the near surface profiles retrieved from the sonde."

- **Line 289:** Water vapor (WV) was introduced before

Corrected

- **Line 304:** The direction seems to have more outlier than the wind speed (3.57% instead of 0.03%). Can this be related to the fact that DAWN is only measuring 5 different positions in forward direction? Would a full conical scan improve the wind direction determination?

I think this is more closely tied to challenges with measuring wind direction at near zero wind speed for both the lidar and the sonde. So it is difficult to assess which wind direction measurement is truly “correct”, thus excluding such points from the statistics was deemed to be a prudent decision.

- **Line 321-324:** Different to Aeolus, DAWN is a coherent wind lidar that "only" measures winds by analyzing the narrow-band backscatter signal from aerosols and clouds. But you use DAWN data also for the validation of Rayleigh winds that are measured in almost aerosol-free atmosphere. Can you still justify your approach?

Our view on validation is that in order to truly assess the accuracy of a space-borne sensor, one needs to compare with data from a reference sensor that is as close to “truth” as possible. As evidenced by experiences from DLR, NASA, NOAA, and other ground-based systems, coherent wind lidar has proven to be the most precise, and highest vertical/spatial resolution airborne wind profiling sensor that we are aware of. DLR has conducted several missions focused on Aeolus Cal/Val with their 2-micron system that are referenced in the paper text. We have done a similar activity here with DAWN. We have limited our comparisons to where sufficient signal was detectable by DAWN so we wouldn’t say that the atmosphere was nearly aerosol-free. So in summary, yes we feel the approach is justified unless somehow molecules move differently than aerosols/clouds. We are not aware of a study that has investigated this aspect.

In terms of “Calibration” of Aeolus Rayleigh signals, DAWN Mie-only measurements are not useful. It is unclear to what extent (or if at all) the presence of significant aerosol would influence the Rayleigh wind retrieval. Such an analysis would require synchronized observations with a coherent and direct detection system like those collected with 2-micron and A2D by DLR.

**Lines 321-324 and 326** How do you perform the projection? How accurate do you know the viewing direction of DAWN? In case you are not pointing to the same azimuth direction, your wind measurements would need to be corrected. Did you consider this issue? Is the laser beam pointing really 90° wrt to the heading angle or to the aircraft reference system? Do you control the angle in case the heading angle changes during the flight leg (e.g. due to changing cross wind conditions), or have you verified how constant the heading was during one Aeolus underflight legs?

The reviewer brings up several good points here. We know DAWN is oriented with the 0 degree scanner position pointed at the nose of the aircraft. We know from the very precise DC-8 INS/GPS data the heading associated with every DAWN wind profile. In cross winds, the reviewer is correct that the aircraft heading required to maintain a true ground relative heading will differ, so an assumption that a 90 degree azimuth stare from DAWN would not exactly match the Aeolus pulse direction/orientation. We account for differences between aircraft heading and 90 degrees when we project vector winds to the Aeolus LOS. But during constant stare operations where we only have an LOS wind speed, we have no ability to reproject the speed to a different angle without some assumptions. We tested the sensitivity of this through assumption of a constant wind

direction profile, which is not particularly valid because there were directional changes as much as 30 degrees. We have amended the text to discuss this point in detail.

“Though the DC-8 flew along the Aeolus laser track where it intersected the 6 km altitude, winds with some component perpendicular to the flight track (i.e. “cross winds”) required the DC-8 to head into the wind to maintain a consistent heading. The difference between the DC-8 and Aeolus heading was taken into account when projecting the DAWN wind vector to the Aeolus view. During the Aeolus underpass on the first flight of the campaign (17-18 April 2019, see Table 1), DAWN was mostly operated in single LOS mode with its beam oriented 90° to the right of the aircraft heading in order to match the sampling of Aeolus. Due to strong cross winds, the DC-8 heading differed by as much as 12° from Aeolus. Based upon intermittent vector wind profiles collected during single LOS operations (shown in Figs. 4a-c), we found that projection to a 102° orientation instead of 90° changed the LOS wind speed by up to 4 m/s. Sensitivity tests assuming a constant wind direction profile across the entire Aeolus underpass showed that correcting the LOS wind speeds from a 90° angle to a 102° orientation resulted in a ~0.2 m/s decrease in Aeolus-DAWN Rayleigh bias but a comparable increase in RMSD. We chose not to incorporate this correction because wind direction was not truly constant throughout the 18 April underpass and the relatively negligible change in validation statistics.”

- **Line 329-331:** Have you tried to use different est. error thresholds in order to verify the sensitivity to these values? This would be an interesting step as other Cal/Val teams reported at the recent Cal/Val workshop that the estimated error calculation seems to vary with time and thus different thresholds might have to be used in different periods. Thus, having the statistical comparison with different thresholds would be interesting.

We agree with you that such comparisons would be interesting if we were dealing with a reprocessed “final” Aeolus dataset. But given the very preliminary nature of the Aeolus data from April 2019 where Laser-A output and signal throughput was far lower than at the beginning of the mission, it is unclear to what extent the expected error parameter is representative of the present day data that has better bias correction and other fixes for various Aeolus measurement/instrument issues. So if we were to find some trends associated with expected error, would they be applicable to present day? We feel it would be best to wait for the reprocessed data to do sensitivity testing.

- **Line 341:** two times “Aeolus”, but the second “Aeolus” should be accumulation...  
Corrected

- **Line 360:** Maybe it would help to introduce all acronyms if not already done before...  
Remaining undefined acronyms for these various Aeolus mission partners are now defined as requested

- **Line 392:** What does UTC observation hour particularly mean? From which event is it counted? It would be better to plot UTC time in the top-x-axes in order to prevent any confusion...

If a flight took off one day and persisted into the following day in terms of UTC time on the U.S. West Coast, as was the case with all flights during the campaign due to the need to underfly Aeolus which passed near California in the 02-03 UTC timeframe, we extended the UTC time axis to values beyond 24 UTC. We recognize your concern and we considered manual

modifications to axis tick marks on all the plots and other ways to explain this. But in the end we felt that the notes in the text such as the following “wind speeds exceeding 50 m/s were present at flight level at the time of the Aeolus overpass near 03 UTC (27 UTC on the cross section time axis).” were sufficient to convey this.

- **Line 404:** With cirrus cloud(s)?

Corrected

- **Line 411:** “excellent agreement” → Can you quantify? E.g., largest deviation, accuracy, precision...

Deriving statistics for one particular profile is not especially meaningful, especially given that there were some time differential between the two observations and drift of the sonde due to high winds. But we feel that most independent observes would agree that the two observations are quite close to each other. But the term “excellent” is in the eye of the beholder, so we have revised the text to say “strong correlation”: “Two sonde releases were coordinated with DAWN vector profiles at ~02.83 to 03.02 (0250-0301) UTC demonstrating strong correlation between sonde and DAWN (Figures 4b-c).”

- **Line 416:** “along a long” → probably correct but sounds strange and funny.

Replaced along with for

- **Line 418:** Sometimes, the UTC times are given with “.” sometimes without.

Including a decimal point depends on the precision required to identify a given feature

- **Line 425:** Full stop missing at the end of the line.

Corrected

- **Line 441:** Startocu → Stratocumulus?

Line ~425 noted that we shortened stratocumulus to stratocu at that point and thereafter

- **Line 459-461:** Have you also analyzed the vertical wind speeds in the vicinity of these mountain waves? Do they lead to additional errors in the Aeolus L2B winds which do not consider vertical winds?

We do not collect vertical motion measurements with DAWN, just vertical profiles of the horizontal winds derived from off-nadir laser pulses

- **Line 514:** 21.6 UTC (UTC is missing)

Corrected

- **Line 517:** was → were (?)

Corrected

- **Line 528-530:** Can you give quantitative numbers here?

Per the comment above, we feel that most observers would agree with our assessment that there is strong agreement between the DAWN profile and sonde, and that reporting statistics on each and every radiosonde comparison in the paper that we show would lessen the readability of an

already relatively long and dense text. We do recognize your concern about the use of terms such as “excellent” and “quite well”, in that quite and excellent are subjective. We have removed the word quite from this sentence, leaving us with “DAWN retrieved a full wind profile that agreed well with a sonde”

- **Line 611-612:** Can you give quantitative numbers here?

We agree with you, however, deriving statistics for one particular profile is not especially meaningful, especially given that there was substantial time and volume differential between the two observations. The following text was amended/added:

“As discussed above, the comparisons with the DLH in-situ open path measurement conducted during a spiral showed good agreement with an average percent difference above 4.5 km and below 1 km (PBL) of approximately 5%. Statistics between 1 km-4.5 km are omitted here due to the sparse sampling statistics and large variability within the in-situ sampling volume. The limited comparison between HALO and DLH show very good agreement and provide confidence in the validity of the measurements throughout the duration of the Aeolus campaign. A HALO WV validation paper is currently in preparation and will provide further details on HALO performance with independent in-situ and space-based observations.”

- **Line 631:** “reached up to 30°” → Probably you would have reached larger values, but you consider differences larger than 30° as gross outlier, right? If so, this should again be mentioned here.

We have revised this portion of the text to reflect uncertainty with wind direction measurement from DAWN and sonde at very low wind speed.

“Wind direction precision decreased with decreasing wind speed and was lowest for wind speed less than 5 m/s (Figure 15c). This is to be expected given that weak wind flows can have variable wind direction over the typical observation/comparison periods discussed here. For example, the sonde wind direction profile deviated from DAWN quite significantly in the 6-7 km altitude layer in Figure 4d where wind speeds less than 2.5 m/s were measured. It is unclear to what extent the sonde can precisely measure wind direction at very slow wind speed, so we feel that use of a 30° gross outlier filter is justified.”

- **Line 679-680:** I would also prefer to read quantitative numbers here instead of “excellent agreement”.

The RMSD values for DAWN have been inserted and the term “excellent agreement” was replaced with “close agreement”