The Optical Autocovariance Wind Lidar. Part II: Green OAWL (GrOAWL) Airborne Performance and Validation

S. BAIDAR

Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, and NOAA/Earth System Research Laboratory/Chemical Sciences Division, Boulder, Colorado

S. C. TUCKER

Ball Aerospace, Boulder, Colorado

M. BEAUBIEN

Yankee Environmental Systems, Turners Falls, Massachusetts

R. M. HARDESTY

Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, and NOAA/Earth System Research Laboratory/Chemical Sciences Division, Boulder, Colorado

(Manuscript received 16 February 2018, in final form 11 July 2018)

ABSTRACT

A two-look airborne Doppler wind lidar operating at the 532-nm laser wavelength, the Green Optical Autocovariance Wind Lidar (GrOAWL), was built and flown aboard the NASA WB-57 research aircraft. Flight campaign goals were to validate the instrument wind measurements and to demonstrate the two-look measurement concept proposed for spaceborne mission concepts such as the Atmospheric Transport, Hurricanes, and Extratropical Numerical Weather Prediction with the Optical Autocovariance Wind Lidar (ATHENA-OAWL) mission. The GrOAWL-measured winds were compared with collocated dropsonde measurements. Line-of-sight velocity (LOSV) measurements for the individual GrOAWL looks showed excellent agreement with dropsondes ($R^2 > 0.9$). The LOSV biases were very small and not statistically different from 0 m s^{-1} at the 95% confidence interval (-0.07 ± 0.07 m s^{-1} and 0.01 ± 0.07 m s^{-1} for look 1 and look 2, respectively). The wind speed and direction profiles retrieved by combining the two GrOAWL looks were also in very good agreement ($R^2 > 0.85$). An instrument performance model indicated the instrument wind measurement precision was likely lowered (uncertainty was increased) by a factor of ~3.3 during the flights relative to predicted "as built" instrument performance. The reduced performance was not observed during ground-based atmospheric testing and thus has been attributed to impacts of the harsh operating conditions of the WB-57 aircraft (high vibration, thermal gradients, and high humidity). The exercise of scaling the GrOAWL instrument performance and grid scale to space showed space-based OAWL wind measurements would yield products with precision at least as good as the GrOAWL instrument.

1. Introduction

Airborne Doppler wind lidars (DWL) have been used for a wide range of applications. Early applications included the study of severe storms (Bilbro et al. 1984) and wind shear detection in front of an aircraft (Targ et al. 1991). Over the years, the use of airborne DWL has expanded to other scientific research areas, such as hurricane forecast improvement (Zhang et al. 2018), gravity wave study (Witschas et al. 2017), long-range transport of pollutants (Chouza et al. 2016), and emissions from urban areas (Baidar et al. 2013), among others. In addition, airborne DWLs also act as prototypes for future space-based systems for global wind profile measurements and are used for instrument testing, validation, and simulation.

Denotes content that is immediately available upon publication as open access.

Corresponding author: Sunil Baidar, sunil.baidar@noaa.gov

DOI: 10.1175/JTECH-D-18-0025.1

© 2018 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

Global wind profile measurement from space is considered as the most important missing observation for improving numerical weather predication (NWP) (WMO 2016a,b; Baker et al. 2014) and the need for such an observation using a DWL and/or passive imaging sensors was once again highlighted in the recently published Decadal Survey for Earth Science and Applications from Space (NASEM 2018). While the performance of any airborne DWL can be scaled for space-based feasibility studies, designated airborne prototypes for spaceborne DWL instruments are needed for instrument-specific performance analysis and retrieval development. The Atmospheric Laser Doppler Instrument (ALADIN) airborne demonstrator (A2D) (Paffrath et al. 2009; Reitebuch et al. 2009) was developed as the airborne demonstrator for the European Space Agency's Aeolus mission (Stoffelen et al. 2005) and it has been extensively used for instrument model validation, wind retrieval testing, and performance analysis of the Aeolus instrument (Lux et al. 2018). It is a single-look, direct-detection instrument operating at 355 nm with a double-edge Fabry–Perot interferometer for molecular winds and a Fizeau interferometer for aerosol winds detection.

The Green Optical Autocovariance Wind Lidar (GrOAWL) was built in 2016 by Ball Aerospace as the prototype instrument for potential space-based missions, including the recently proposed Atmospheric Transport, Hurricanes, and Extratropical Numerical Weather Prediction with the Optical Autocovariance Wind Lidar (ATHENA-OAWL) mission for the NASA Earth Venture Instrument (EVI) program (Weimer et al. 2015; Tucker et al. 2016). The GrOAWL instrument is a direct-detection system based on the Optical Autocovariance Wind Lidar (OAWL) concept employing a quadrature Mach–Zehnder interferometer (QMZI) receiver (Grund et al. 2009). It operates at 532-nm wavelength and has two azimuthally orthogonal lines of sight (LOS) to measure two components of the horizontal winds from aerosol returns. While single-LOS measurements from space like the Aeolus mission can provide up to 70% of the vector wind information under optimal conditions (Horányi et al. 2015), the two azimuthally orthogonal LOS measurements provide horizontal wind speed and direction information.

In this study, the GrOAWL instrument performance during the ATHENA-OAWL Venture Tech (AOVT) airborne flights aboard the NASA WB-57 aircraft in 2016 is presented. The emphasis of the airborne deployment was on using independent wind measurements to demonstrate the GrOAWL's ability to measure winds from an airborne platform. The GrOAWL-measured LOS velocities (LOSVs) for each look were individually validated using collocated dropsonde wind measurements. Wind speed and direction retrieved by combining data from the two looks were also compared with the dropsonde measurements. In-flight instrument performance was evaluated using instrument performance model simulations. Simulations of the satelliteborne OAWL instruments were also performed and evaluated against the GrOAWL airborne performance to demonstrate the feasibility for space-based OAWL wind measurements. A follow-up paper will describe the ground-based study focused on verifying instrument performance and validating the instrument performance model in a well-characterized atmosphere.

2. Measurements

a. *AOVT flight campaign*

The GrOAWL instrument was installed on a NASA WB-57 research aircraft for the AOVT flight campaign. The flights took place over the Gulf of Mexico off the Texas coast between 10 May and 24 June 2016. The main objectives of the flight campaign were to (i) validate the GrOAWL winds with independent measurements and (ii) demonstrate the two-look ATHENA-OAWL mission concept to measure two independent LOS wind components and to demonstrate combining the two looks to obtain horizontal wind speed and direction estimates. The campaign consisted of a total of eight research flights on the NASA WB-57 research aircraft. The first five flights were considered engineering flights with only the GrOAWL instrument aboard the aircraft. During the last three flights, an automated dropsonde system was added to the WB-57 payload bay for validating the GrOAWL wind measurements. Radiosondes launched from the Corpus Christi, Texas, National Weather Service (NWS) office during the validation flights provided additional data. Because of restrictive laser eye safety regulations on the 532-nm wavelength for other aircraft, all flights were performed over designated aircraft warning areas over the Gulf of Mexico. Flying over warning airspace provided flexibility to launch dropsondes for comparison and also to make modifications to the flight plan to avoid convective cells in the flight path. Figure 1a shows the flight paths for the three validation flights. All flights originated from Ellington Field, Houston, Texas, and consisted of racetrack patterns lasting around 60–90 min per loop at altitudes between 8.5 and 12 km (~12–17-km range to the ocean surface). The racetrack patterns were designed such that the same atmospheric volume enclosed by the racetrack was interrogated by the GrOAWL LOS while flying the two legs of the racetrack, capturing atmospheric variability over the course of the flight. During each leg, three to four dropsondes, with equal horizontal spacing,

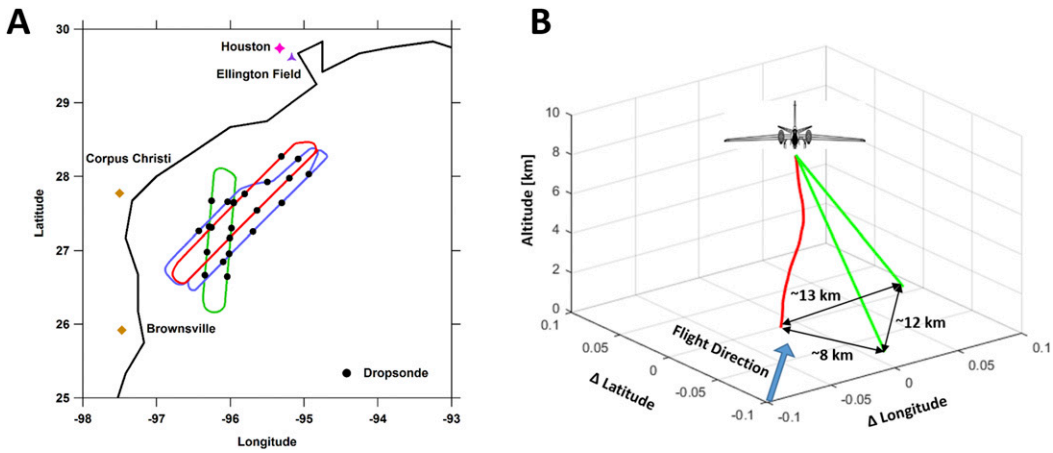

FIG. 1. (a) NASA WB-57 research aircraft flight paths during the three AOVT validation flights with the GrOAWL system (red: 17 Jun; green: 21 Jun; blue: 24 Jun). The black closed circles show locations of the dropsonde launches from the aircraft. All flights originated from Ellington Field. (b) Schematic showing the two GrOAWL lines of sight (green) and a dropsonde path (red). The distances represent the distance at the surface between (i) dropsonde and the closest GrOAWL measurements and (ii) separation between two GrOAWL looks at the surface.

were launched over the same geographical positions during the multiple loops of the flight. Figure 1a shows the geographical positions of the dropsonde releases as black closed circles.

b. *GrOAWL instrument*

The GrOAWL instrument is described in detail in Tucker et al. (2018). Briefly, the GrOAWL is a direct-detection DWL measuring winds from aerosol backscatter at 532 nm using optical autocovariance techniques first introduced by Schwiesow and Mayor (1995). It measures the Doppler shifts of backscattered light relative to the outgoing laser pulses as phase shifts in interferometer fringe function using a field-widened, four-channel Mach–Zehnder interferometer (MZI) (Liu and Kobayashi 1996; Bruneau and Pelon 2003; Grund et al. 2009). The OAWL implementation is a self-referencing system: only the shift in the phase is important for wind measurement and the Doppler-shifted atmospheric returns are referenced to the phase of the corresponding outgoing time-zero (T0) pulse, which is continuously monitored.

Like the ATHENA-OAWL instrument design, the GrOAWL instrument includes two lasers and two telescopes for two looks coupled into one interferometer receiver. For GrOAWL, the telescopes are 28-cm effective diameter pointing 45° off nadir and at 45° and 135° azimuths relative to the aircraft longitudinal axis. Two neodymium-doped yttrium aluminum garnet (Nd:YAG) lasers send pulses alternatively at 200 Hz each at both 532 and 355 nm, and the backscattered light is recorded at a raw 140-MHz (1.07 m) rate on all four interferometer channels. The atmospheric returns are binned to 8.56-m range gate resolution (~5-m vertical resolution) to reduce the data volume. Both 532- and 355-nm transmitters/receivers were on board because of the High-Spectral-Resolution Lidar (HSRL) for Aerosols, Winds, and Clouds Using OAWL (HAWC-OAWL) program (Tucker et al. 2018), but only the 532-nm wavelength data are presented here, as the AOVT campaign was aimed at demonstrating OAWL operation at 532 nm. The instrument was operated at ≤1.5 mJ at 532 nm during the AOVT flights. Table 1 lists the GrOAWL system specifications as deployed and processed for the AOVT campaign.

c. *Dropsonde*

The High-Definition Sounding System (HDSS) from Yankee Environmental Systems (YES) was installed immediately aft of the GrOAWL instrument in the NASA WB-57 aircraft for the validation flights. The HDSS is an automated system that deploys the expendable digital dropsonde (XDD) to measure wind, pressure, temperature, and humidity profiles (Black et al. 2017). Unlike the conventional dropsondes, such as Vaisala RD94, the XDD does not have a parachute. It can measure winds at a 4-Hz sampling rate (~5–8-m vertical resolution depending on dropsonde fall rate) with an accuracy similar to that of conventional dropsondes (0.5 m s^{-1}). A total of 48 HDSS dropsondes were launched from the WB-57 during the three validation flights and 44 of the dropsondes produced valid results. It took ~6 min for the dropsondes to reach the

TABLE 1. System specifications of the GrOAWL instrument during AOVT flights and SpOAWL instruments used for simulations. SpOAWL instruments differ only in the telescope size. The telescope diameters are 70, 100, and 150 cm for SpOAWL1, SpOAWL2, and SpOAWL3, respectively. SpOAWL1 represents the EVI-2 ATHENA-OAWL instrument configuration. ISS is International Space Station.

Instrument	GrOAWL	SpOAWLs
Platform	NASA WB-57	ISS or free flyer
Nd:YAG wavelength	532 nm	532 nm
Laser pulse energy	≤1.5 mJ	160 mJ
Pulse repetition frequency	200 Hz	150 Hz
MZI OPD	0.9 m	0.9 m
ITE	1.0, 0.6	1.0
IC	0.73/0.33	0.9
Off-nadir point angle	45°	40°
Azimuth angles (from forward)	45° and 135°	45° and 135°
Effective telescope diameter	28 cm	70/100/150 cm
Number of telescopes	2	2
Flight altitude	8–12 km	400 km
Range	Ground to 0.5 km below flight level	0–25 km
Sampled data resolution	8.56 m at 200 Hz	30 m at 150 Hz
Retrieved wind resolution (notional)	~121 m × 10 s	1 km × 12 s
N	40 000	77 400

surface from ~8.5-km flight altitude and ~8 min from ~11.5-km flight altitude.

3. Results

a. LOSV

The GrOAWL LOSV retrieval process is described in Tucker et al. (2018). First, the phase of every outgoing pulse (T0 phase) is determined by fitting a sinusoid to the signal at each of the four detectors. The aircraft-platform-motion-induced Doppler (phase) shift [calculated using on-instrument GPS-internal navigation unit (INU) data] is then added to the T0 phase. This platform-motion-corrected T0 phase is used as the reference phase for the atmospheric returns on a per-pulse basis. The reference-phase-corrected atmospheric returns are then accumulated over range gates and time (laser pulses). A sinusoidal fit is performed to the accumulated data at each of the four detectors to retrieve atmospheric-wind-induced phase shift. The relationship between LOSV v_{LOS} and phase shift $\Delta\phi$ is given by

$$v_{\mathrm{LOS}} = \frac{\lambda c}{4\pi \mathrm{OPD}} \Delta\phi. \tag{1}$$

Because the instrument records and stores data at very high vertical and temporal resolution, it is possible to process the GrOAWL data at different resolutions depending on the scientific need or backscattered signal strength. For example, inside the boundary layer, where the winds are highly variable and aerosols are more prevalent, it would be scientifically relevant to retrieve winds with high vertical and temporal resolution. In contrast, in the upper troposphere, where aerosol loading is low, it might be necessary to accumulate a larger number of pulses and/or range gates to tease out the signal. For the purpose of this paper, LOSV profiles were retrieved at 10-s time resolution and ~121-m vertical resolution (i.e., 2000 laser pulses and 20 sample range gates). This temporal and vertical resolution was chosen to provide LOSVs with a precision better than 2 m s^{-1} for most of the vertical profiles (see section 4 for details), and this level of uncertainty was determined to be sufficient to establish the GrOAWL's ability to accurately measure winds during the flights. A 10-s temporal resolution corresponds to about a 1300-m spatial resolution at the nominal aircraft speed of 130 m s^{-1}. Figure 2 (top panels) shows a time–height cross section of LOSV profiles for the two GrOAWL looks from one racetrack loop on 17 June 2016 (Fig. 1, red track). The aircraft turned at 1800 and 1845 UTC. Depending upon the flight orientation, for a given GrOAWL look the winds (in Earth frame of reference) may be blowing toward the system on one leg of the flight loop and thus away from it on the return leg. This results in the LOSVs for the same look having opposite signs (e.g., cool vs warm colors) on either side of the turn as seen in Fig. 2 (top panels). The corresponding measurement carrier-to-noise ratio (CNR) profiles are shown in the Fig. 2 (bottom panels). CNR is a function of the lidar signal-to-noise ratio (SNR) (number of lidar return photons) and measurement contrast (a function of atmospheric scatting ratio and interferometer alignment-based instrument contrast). Instrument contrast is range independent for a given profile and hence variability in CNR with altitude is

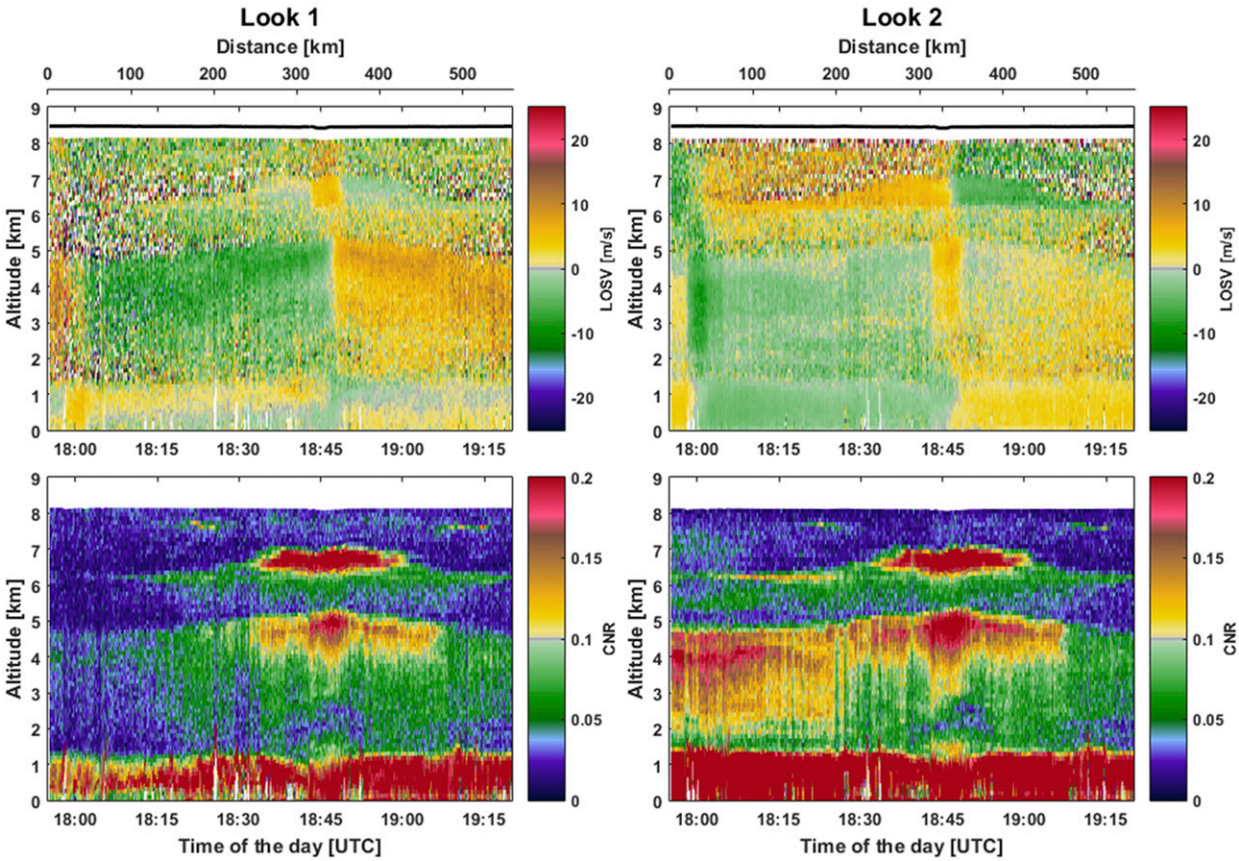

FIG. 2. (top) Time–height curtain plot showing the GrOAWL LOSV profiles during one loop of the racetrack pattern flown on 17 Jun 2016 (Fig. 1a, red track) for the two GrOAWL looks. The temporal and vertical resolutions of the data are 10 s and ~121 m, respectively. The aircraft turned at 1800 and 1845 UTC, prompting the change in the measured LOSV sign. The black horizontal line indicates the flight altitude. The top axis shows flight distance. (bottom) Time–height curtain plot showing the corresponding measurement CNR. As both looks use the same interferometer and detector, the differences in CNR between the two looks are mainly attributed to differences in the quality of the telescope alignment/overlap, which impacts the lidar SNR.

due to variability in lidar SNR and aerosol content with altitude. As both looks use the same interferometer and detector, differences in CNR between the two looks are mainly due to differences in the quality of the telescope alignment/overlap, which impacts the lidar SNR.

Figure 3 shows an example of GrOAWL and dropsonde LOSV profiles from the flight on 17 June 2016. Dropsonde LOSV profiles were calculated from dropsonde-measured horizontal wind vector (u and v) using

$$v_{\mathrm{LOS}_i} = u \sin\theta_i \cos\alpha_i + v \cos\theta_i \cos\alpha_i \,, \qquad (2)$$

where θ_i and α_i are aircraft attitude-corrected azimuth and elevation angle for look i, respectively, in the Earth coordinate system.

The GrOAWL LOSV profiles in Fig. 3 represent an average of six retrieved profiles, (a 1-min average, less than the time for the sonde to reach the surface). Because of the difference between the GrOAWL LOS pointing and the dropsonde fall path, there is no direct spatial overlap between the two measurements. Spatial mismatches between dropsonde tracks and the GrOAWL lines of sight are illustrated in Fig. 1b. The dropsonde shown in Fig. 3 was launched at 1856 UTC and reached the ocean surface at 1902 UTC. Thus, the intercomparison shows data up to 8 km apart in space and 6 min in time. The gray shading represents the GrOAWL LOSV 1σ standard deviation over a 6-min time period, and the pink shading is an estimate of the upper limit for instrument-induced error determined assuming the minimum of the GrOAWL LOSV 1σ standard deviation is entirely due to the instrument noise. The agreement between the two types of measurements is very good at all altitudes with the GrOAWL measurements capturing all the features observed by the dropsonde. This very good agreement was typical for other dropsondes during the AOVT flights [see section 3a(2)].

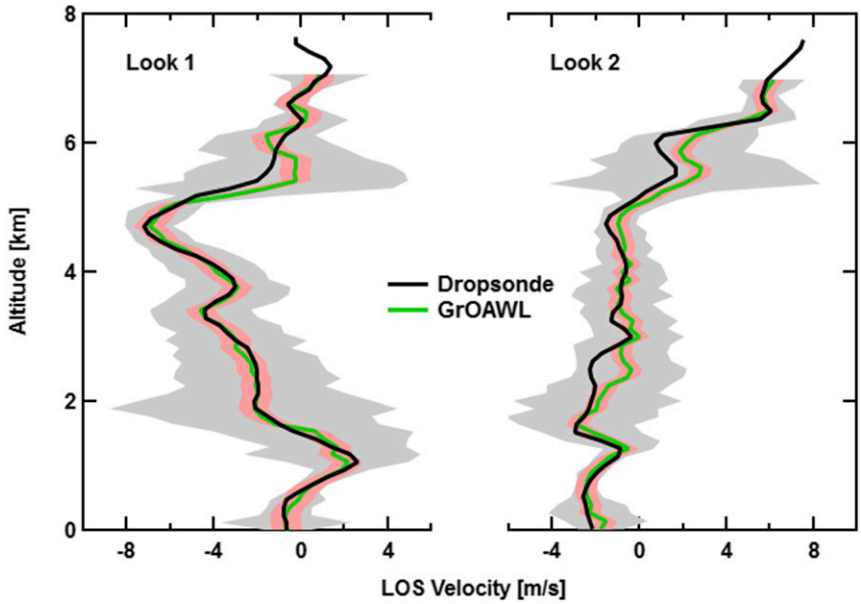

FIG. 3. GrOAWL LOSV profiles (green) for the two looks corresponding to a dropsonde (black) launched during the 17 Jun 2016 flight. The GrOAWL profiles are 1-min averages of 10-s profiles, and the gray shading depicts a 1σ standard deviation over a 6-min period, which is the time the dropsonde took to reach the surface. The pink shading, which is the minimum of the gray shading over the altitude range, represents the upper limit for instrument-induced error.

1) LOSV PRECISION

The GrOAWL LOSV measurement precision, like that for all aerosol wind lidars, depends upon instrument parameters included in the basic elastic lidar equation (e.g., pulse energy, telescope size, system efficiency), instrument sensitivity (interferometer alignment and laser coherence length), atmospheric parameters (e.g., aerosol backscatter coefficient), and processing parameters (e.g., range gate size, integration time). The Cramer–Rao lower bound (CRLB) on the precision of GrOAWL LOSV is a function of CNR and the number of samples N in the accumulation and is given by

$$\sigma_{\mathrm{LOSV}}(N) \geq \frac{\sqrt{2}}{\mathrm{CNR}} \frac{c\lambda}{4\pi \mathrm{OPD}} \frac{1}{\sqrt{N}}, \tag{3}$$

where OPD is the optical path difference of the interferometer (Tucker et al. 2018).

The CNR-based GrOAWL precision profiles calculated using Eq. (3) for every 10-s accumulation LOSV profile for look 2 on 17 June 2016 flight are shown in Fig. 4. The black line is the median of the precision distribution, and the gray shaded region represents the 10th and 90th percentiles of the distribution. Variability in the CNR along the flight track (Fig. 2, bottom panels) is the cause for the precision distribution. The LOSV precision is, on average, $\sim 0.5\ \mathrm{m\,s^{-1}}$ in the boundary layer, where there are abundant aerosols. In the free troposphere, precision estimates show both degraded precision and wider distribution, indicating a relatively lower average concentration and larger variability in aerosol. Instrument-induced error is range independent, whereas uncertainty

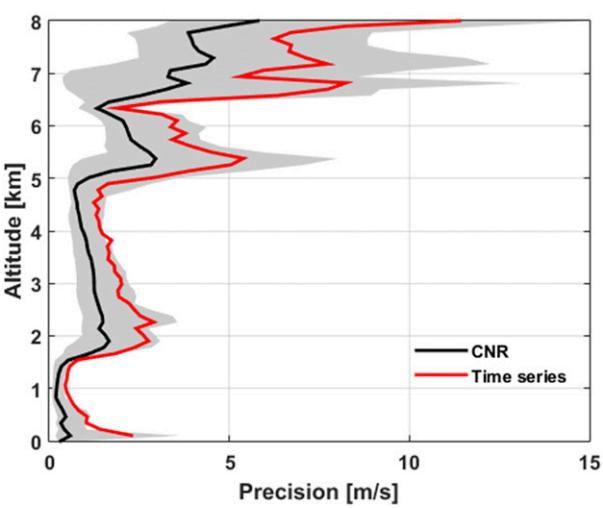

FIG. 4. GrOAWL LOSV precision for look 2 during the 17 Jun 2016 flight based on instrument CNR (black) and detrended LOSV variability (red) in each range gate. The black line is the 50th percentile of the precision distribution, and the gray shading represents the 10th and 90th percentiles of the data. The precision estimates are for 10-s profile data at ~121-m vertical resolution.

because of aerosol content, which also affects measurement CNR, can vary with range. The best-achieved average LOSV precision for these processing parameters is ~0.3 m s^{-1} at ~1-km altitude. This value represents an upper bound on instrument-induced uncertainty (see pink shading in Fig. 3). Any reduction in LOSV precision (and corresponding increase in LOSV precision distribution) with altitude is a result of lower aerosol content (and higher aerosol variability). The observed degraded precision with altitude caused by lower aerosol content is as expected based on the results from the instrument performance model using the Goddard Earth Observing System, version 5 (GEOS-5), aerosol profiles (see section 4 for details). Note that a decrease/reduction in measurement precision refers to an increase in measurement uncertainty and vice versa.

Precision of a measurement that varies with time or location can be defined as the standard deviation of variability in detrended time series data (e.g., Gisi et al. 2012). Average GrOAWL LOSV precision at each altitude is calculated from detrended measured LOSV time series at that altitude. This can be used to validate the CNR-based precision estimates. The red line in Fig. 4 shows the precision (1σ standard deviation) determined from the LOSV variability for each range gate. Atmosphere-induced trends in the LOSV were removed by subtracting a five-point (e.g., 50 s) moving average. Atmospheric variability/motions that are smaller in scale than the detrending time period (e.g., 50 s here) are not entirely removed by the detrending process and hence the calculated average LOSV precision is generally larger that the true LOSV measurement precision. Further, 10-s accumulation LOSV profiles do not resolve smaller-scale motions. Thus, calculated average LOSV precision values are expected to be larger than the median CNR-based LOSV precision estimates, based on the theoretical lower bounds. Nevertheless, the very good agreement (r = 0.90) between the two precision estimates provides the closure between OAWL theory [Eq. (3)] and measurements, and these two precision profiles provide bounds for the true LOSV measurement precision.

2) LOSV VALIDATION AND ACCURACY

While the GrOAWL wind measurement precision depends on instrument, atmospheric, and processing parameters, and can be improved or degraded by changing these parameters, accuracy of the measured winds is inherent to the instrument design, processing algorithms, and knowledge of platform velocity and pointing. The accuracy of the GrOAWL-measured LOSVs for each look was evaluated using LOSVs calculated from collocated dropsonde wind measurements. In this analysis, accuracy of the measurement is defined as bias relative to a reference measurement. Figure 5 shows scatterplots comparing LOSVs between the GrOAWL and dropsonde for two looks. The scatterplots include data from a total of 21 dropsondes from three validation flights ($N > 800$). Because of instrument issues (mainly thermal and vibrational), measurements from the latter half of the flights were not included in the comparison. Dropsonde-measured winds were projected onto the GrOAWL LOS, interpolated to the center of the GrOAWL measurement grid, and compared to the closest GrOAWL measurements in location. The GrOAWL data represent 10-s accumulated profiles. In Fig. 5, data filtered for different GrOAWL measurement precision thresholds [calculated using Eq. (3)] are plotted on top of each other. The lower-precision (higher uncertainty) thresholded data are plotted first and then successively higher-precision (lower uncertainty) thresholded data are stacked on top using increasingly darker markers. The red line is the linear fit to the 2 m s^{-1} precision filtered data. Precision of the data is important for setting boundaries on the uncertainty in the assessed accuracy. Accuracy of measurement is usually determined at high SNR. Hence, the accuracy of the measurement was assessed at different precision levels. The results from the linear regression fits to the data filtered for different precision thresholds are given in Table 2. The uncertainties reported in Table 2 represent the 95% confidence interval.

As seen in Table 2, tighter constraints on the precision give more robust results with R^2 values greater than 0.85. However, these results are based on a limited subset of data (~20% of the data are within 1 m s^{-1} precision threshold; N = 238 for look 1), most of which are from inside the boundary layer, where there are abundant aerosols for signal (see Fig. 4). Thus, results from the tighter-precision thresholds might not be representative of the entire atmosphere. Vertical coverage can be improved by loosening the precision threshold (~65% of the data are within 4 m s^{-1} threshold); however, this increases scatter and decreases the correlation coefficient. Nevertheless, the slopes of the linear fit are within ±12% of unity, across the different precision thresholds for both looks, indicating very good agreement between the GrOAWL and dropsonde measurements. The intercept ranged from −0.13 to 0.21 m s^{-1} and from 0.10 to 0.29 m s^{-1} for look 1 and look 2, respectively. Uncertainty in the intercept of the linear fits is around 0.2 m s^{-1} and shows negligible dependence on the different precision thresholds, because while lower thresholds reduce scatter, they also result in fewer points for the linear fit, while higher thresholds provide more points but also result in larger scatter.

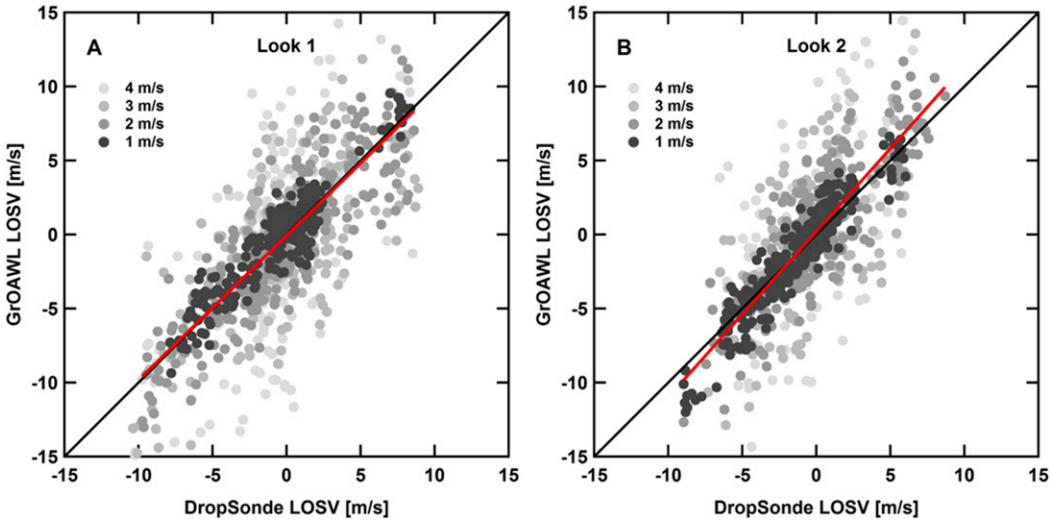

FIG. 5. Scatterplot comparing LOSVs from the two GrOAWL looks and dropsondes. A total of 21 dropsondes from the three validation flights are included in the comparison. Dropsonde LOSVs are projections of dropsonde-measured wind speeds and directions. The GrOAWL data are 10-s profiles that are spatially closest to the dropsonde measurement. Data points are stacked by the GrOAWL measurement precision. The red line is the linear regression fit to the 2 m s^{-1} precision threshold data, and the black line is the 1:1 line. Results from the regression fit for different precision thresholds are given in Table 2.

GrOAWL's performance degrades gradually with CNR, such that precision of the measurements at low CNR regions can be improved by accumulating more pulses or samples in the retrieval. Hence, useful data can be retrieved even at low CNR. Note that although accumulation generally increases the measurement precision, it does not improve nor degrade measurement accuracy. For stationary wind fields, averaging retrieved profiles can also improve wind precision, without degrading accuracy. Averaging retrieved profiles is more practical than reprocessing to accumulate additional pulses or along-range samples to meet certain precision thresholds, assuming the errors are uncorrelated, as the required number of pulses and samples for accumulation will vary by range and aerosol conditions.

In this analysis, the number of higher-precision GrOAWL data points in the dropsonde comparison is increased by averaging 10-s profiles in order to obtain a more robust result for the measurement accuracy assessment. Figure 6 shows scatterplots comparing 2-min-averaged GrOAWL data (12 profiles) and dropsonde observations for two looks. The averaged GrOAWL data were thresholded for average precision above 4 m s^{-1} so that averaging over 2 min effectively made the precision of all the data points (standard deviation) in the scatterplot better than 2 m s^{-1}, assuming the errors are uncorrelated. Indeed, the RMSEs of 1.04 and 0.98 m s^{-1} for look 1 and look 2, respectively, are comparable to RMSEs for the 1 m s^{-1} precision-filtered 10-s data (see Table 2). Further averaging the GrOAWL data also did not change the slope and intercept of the linear regression (not shown). This indicates that the atmosphere was stable during the time of the measurements, providing ideal conditions for dropsonde comparison and demonstrating that the dropsonde measurements were representative of the true mean wind conditions.

Averaging the GrOAWL data provides a comparison of the mean wind measurements within a certain spatial volume with the dropsondes. The 2-min time period represents a 16-km spatial averaging at the WB-57 nominal speed. The results of the linear regression fits to the data with free slope (red lines in Fig. 6) and fixed slope (set to 1) are given in Table 3. The data of both looks show very good agreement with R^2 values greater than 0.90. The linear regression fits with fixed slope equal to 1 indicate small but statistically insignificant biases of -0.07 ± 0.07 m s^{-1} and 0.01 ± 0.06 m s^{-1} (95% confidence interval) for look 1 and look 2, respectively. Thus, the GrOAWL wind accuracy is similar to that of the dropsondes. Since the GrOAWL is a self-referencing system, these small biases are very likely related to the aircraft motion correction. Improper aircraft motion correction, including incorrect telescope pointing angle knowledge, introduces bias in the measured LOSV. Uncertainty in the aircraft motion and telescope pointing angle resulting from jitter increases the uncertainty

TABLE 2. Results from linear regression with dropsonde measurements of the GrOAWL 10-s data filtered for different precision thresholds. The reported uncertainty represents the 95% confidence interval. Bias is calculated as the mean difference between GrOAWL and dropsonde measurements. The uncertainties in the bias represent the standard error of the mean.

		Precision			
		$<1\,\mathrm{m\,s^{-1}}$	$<2\,\mathrm{m\,s^{-1}}$	$<3\,\mathrm{m\,s^{-1}}$	$<4\,\mathrm{m\,s^{-1}}$
Look 1	N	238	499	724	865
	Slope	0.98 ± 0.05	0.98 ± 0.04	0.94 ± 0.05	0.96 ± 0.06
	Intercept ($\mathrm{m\,s^{-1}}$)	0.21 ± 0.16	−0.05 ± 0.17	−0.10 ± 0.19	−0.13 ± 0.22
	Bias ($\mathrm{m\,s^{-1}}$)	0.22 ± 0.08	−0.04 ± 0.09	−0.08 ± 0.10	−0.11 ± 0.11
	RMSE ($\mathrm{m\,s^{-1}}$)	1.24	1.96	2.61	3.26
	R^2	0.87	0.79	0.67	0.57
Look 2	N	304	600	749	838
	Slope	1.12 ± 0.04	1.12 ± 0.05	1.08 ± 0.05	1.08 ± 0.06
	Intercept ($\mathrm{m\,s^{-1}}$)	0.10 ± 0.14	0.22 ± 0.15	0.20 ± 0.17	0.29 ± 0.20
	Bias ($\mathrm{m\,s^{-1}}$)	−0.07 ± 0.07	0.14 ± 0.08	0.15 ± 0.09	0.24 ± 0.10
	RMSE ($\mathrm{m\,s^{-1}}$)	1.21	1.85	2.41	2.87
	R^2	0.89	0.80	0.68	0.60

of the measured LOSV. The precision of the GPS–inertial measurement unit (IMU) system is $0.03\,\mathrm{m\,s^{-1}}$, 0.005°, and 0.03° for velocity, pitch and roll angles, and heading, respectively. Thus, the uncertainty resulting from aircraft motion correction and/or pointing angle is expected to be constant with range and negligible compared to the atmospheric variability. The excellent agreement shown in Figs. 5 and 6 indicates the fundamentally correct operation of the GrOAWL instrument and the wind retrieval algorithm.

b. *Wind speed and direction retrieval*

The two GrOAWL looks are azimuthally orthogonal. Hence, assuming the average vertical wind velocity is negligible (or $0\,\mathrm{m\,s^{-1}}$) over the time scales of the GrOAWL measurements, horizontal components of the LOSVs for the two looks represent the two components of the horizontal wind vector (u' and v') and are given by

$$u' = \frac{v_{\mathrm{LOS}_1}}{\cos\alpha_1} \quad \text{and} \quad v' = \frac{v_{\mathrm{LOS}_2}}{\cos\alpha_2}, \tag{4}$$

where α_i is the elevation angle for look i in the Earth coordinate system, which is a function of aircraft orientation and instrument-to-aircraft alignment. Except in strong convective conditions, vertical velocity tends to be very small and difficult to measure (an order of a few centimeters per second; Holton 2004) and hence can be considered to have negligible contribution to LOSVs. For example, average vertical velocity for the flight track loop shown in Fig. 2 was $0.01\,\mathrm{m\,s^{-1}}$. Vertical velocity was calculated from dropsonde measurements using the method described in Wang et al. (2009).

Horizontal wind speed and direction are calculated from these pseudo-u and -v components of wind vector using Eqs. (5) and (6), respectively:

$$\text{Wind Speed} = (u'^2 + v'^2)^{1/2}, \tag{5}$$

$$\text{Wind Direction} = \mathrm{atan2}(v', u') + 180° + \theta_1 + \omega, \tag{6}$$

where θ_1 is the look 1 azimuth angle, relative to the aircraft frame of reference; and ω is the aircraft heading.

Once the LOSVs for the two looks were validated, data from the two looks were combined to retrieve horizontal wind speed and direction using Eqs. (5) and (6). Before calculating wind speed and direction, LOSVs for each look were interpolated to a common vertical grid (200 m in this case) to facilitate combination. The interpolation to a common grid was necessitated by deviations of aircraft attitude (e.g., pitch angle) that resulted in the two GrOAWL looks having a different number of range gates between the aircraft and the ground surface. Furthermore, the LOSVs were averaged for 2 min to ensure that the two looks overlapped at the ground and a single time period can be assigned to a profile. While flying at 8.5-km altitude, the two looks are separated by ~12 km at the surface, which is about 1.5 min of flight time at the WB-57 nominal flight speed (see Fig. 1b). This averaging was done for simplicity and also to facilitate comparison with the dropsonde measurements.

Figure 7a shows a wind-barb curtain plot of the horizontal wind vector for the 17 June 2016 flight. Data from one loop of the racetrack is shown in the figure. It shows temporal and spatial variability of the wind field within the loop. The wind was blowing from the south and fairly light below 1500 m. Between 1500 and 5000 m, the wind

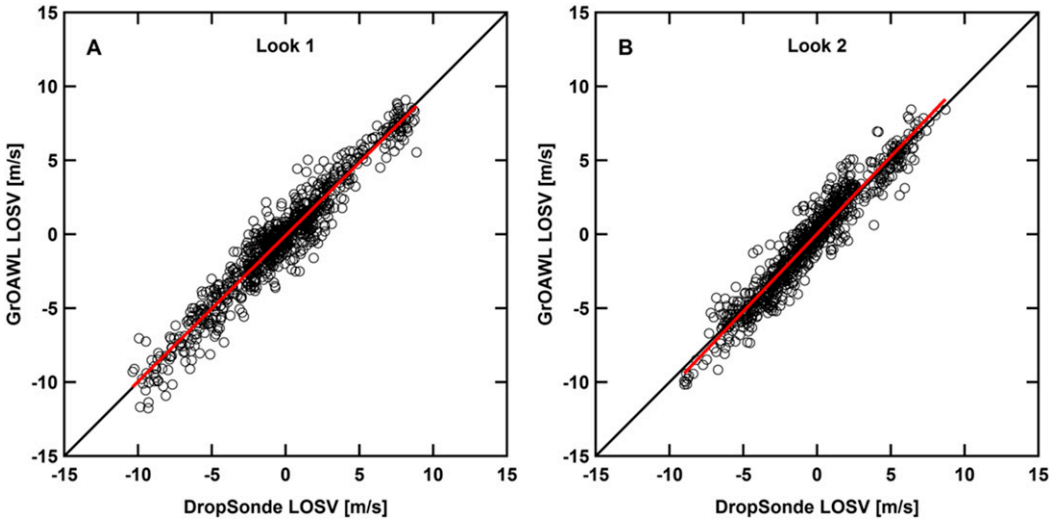

FIG. 6. As in Fig. 6, but with time-averaged GrOAWL data. The GrOAWL data points represent 2-min averages of 10-s profile data around the spatially closest measurement to the dropsonde data (N = 873 and 881 for look 1 and look 2, respectively). A 4 m s^{-1} precision threshold was applied to the lidar data. The red line is the linear regression fit to the data, and the black line is the 1:1 line. Results from the regression fits are given in Table 3.

shifted to easterly. The shift in wind direction at ~600 m likely represents the planetary boundary layer height. There is a layer between 5000 and 6000 m with very low wind speed (<1 m s^{-1}), where determination of the wind direction is difficult. More importantly, the results show that the GrOAWL is able to accurately measure very low wind speeds. The aircraft was flying at ~45° and ~225° heading (Earth coordinates) during the two straight legs (see Fig. 1a, red track). This results in look 1 being oriented in an east–west axis and perpendicular to the predominant wind direction below 1500 m. Similarly, look 2 is oriented in the north–south axis and perpendicular to the wind direction between 1500 and 5000 m. The perpendicular orientation of the telescopes to the wind direction results in very small LOSVs (see Fig. 2). The wind speed and direction profile at the turn (around 1845 UTC) was retrieved using the velocity–azimuth display (VAD) method (Browning and Wexler 1968) and shows very good agreement with the two look profiles on either side of the turn. Six dropsonde wind profiles from the same flight loop are displayed in Fig. 7a in bold. Despite the spatial and temporal separation, the GrOAWL and dropsonde profiles show excellent correspondence. The wind flow over the flight region is better visualized in the 3D plot shown in Fig. 7b.

WIND SPEED AND DIRECTION COMPARISON

Figure 8 shows scatterplots of the GrOAWL wind speed and direction against the corresponding dropsonde measurements. The GrOAWL profile closest in time to the dropsonde launch time is considered for comparison. Results were virtually identical when the GrOAWL wind speed and direction estimates closest in space to the dropsonde measurements were used. The red line represents the linear regression fits to the data, and the black diagonal line is the 1:1 line. Fits for both wind speed and direction give slopes close to 1 and offsets close to 0. Results from the linear regression are given in Table 3. Constraining the slope to 1 in the linear regression results in a bias of 0.28 ± 0.12 m s^{-1} and 4.5° ± 1.6° for wind speed and direction, respectively. This excellent agreement between the two datasets is not surprising, considering the excellent agreement in individual LOSVs between the GrOAWL and dropsondes

TABLE 3. Results from linear regression of the GrOAWL 2-min averages (12 profiles) with dropsonde measurements. The reported uncertainties represent the 95% confidence interval.

	N	Intercept (m s^{-1})	Slope	Intercept for fixed slope (m s^{-1})	RMSE (m s^{-1})	R^2
Look 1	873	−0.07 ± 0.07	0.99 ± 0.02	−0.07 ± 0.07	1.04	0.93
Look 2	881	0.03 ± 0.06	1.05 ± 0.02	0.01 ± 0.07	0.98	0.92
Wind speed	521	0.44 ± 0.26	0.97 ± 0.03	0.28 ± 0.12	1.38	0.86
Wind direction	521	10.4 ± 3.3	0.95 ± 0.03	4.5 ± 1.6	18.61	0.90

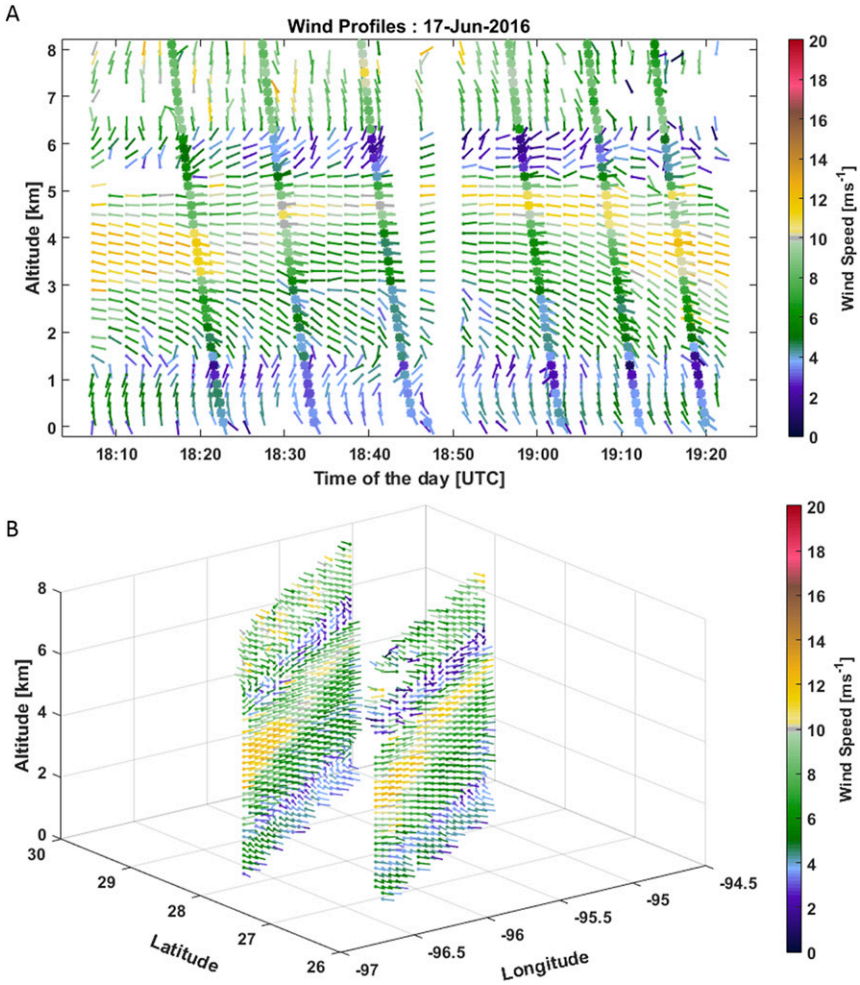

FIG. 7. (a) Wind speed and direction profiles measured by the GrOAWL during the 17 Jun 2016 flight. Wind speed and direction profiles were retrieved by combining LOSVs from the two azimuthally orthogonal GrOAWL looks. Vertical velocity is considered to be negligible in the retrieval. The profile at the turn (around 1845 UTC) was retrieved using the VAD method (Browning and Wexler 1968). Profiles from six dropsondes launched during the same flight loop are shown in bold. (b) A 3D plot of the same data. Northern leg of the flight loop is shifted by 1.25° latitude for display purposes.

[see section 3a(2)]. Although the dual-look space-based wind lidar concepts specify assimilation of the LOSVs directly into NWP models, the excellent agreement here indicates that the two orthogonal components are sufficient to produce good estimates of the horizontal wind vectors under the conditions encountered during the flights.

4. Instrument performance

The GrOAWL instrument performance during the flights was impacted by hardware issues related to the operating environment of the WB-57 research aircraft. Large temperature variations caused changes in telescope focus and lidar beam overlap, and high Houston humidity levels resulted in some window fogging at higher altitudes. High aircraft vibration negatively impacted telescope and some interferometer alignments. These operating-environment-related engineering challenges and subsequent instrument modifications are discussed in Tucker et al. (2018). The overall impact of these instrumental issues on the wind measurement precision is described here.

A radiometric model for the GrOAWL instrument was developed to assess the performance of the instrument during the AOVT flights. The model includes the GrOAWL instrument (laser transmitter, receiver, detector, and alignment parameters) and the atmosphere

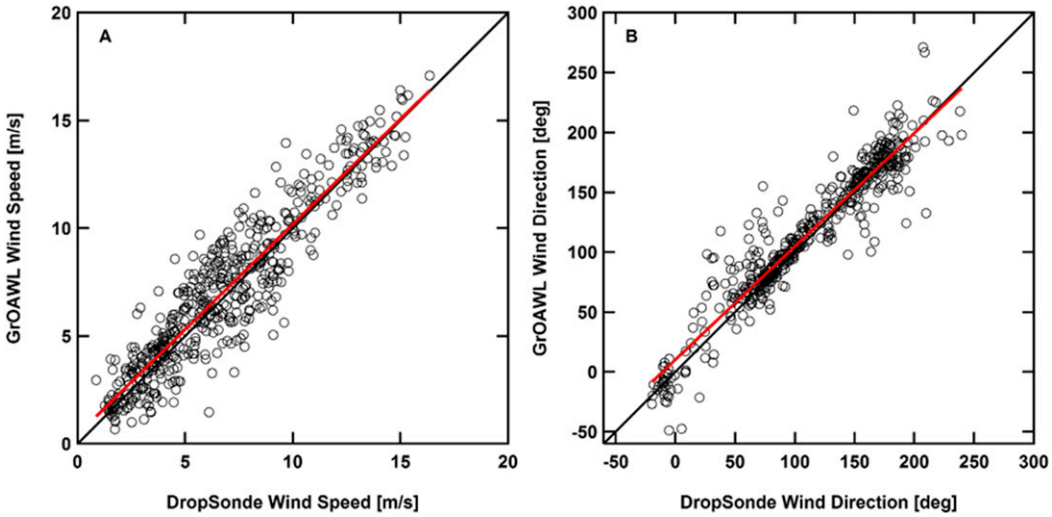

FIG. 8. Scatterplot comparing (a) horizontal wind speed and (b) wind direction measured by the GrOAWL and dropsondes ($N = 521$) during the AOVT validation flights. The red line is the linear regression fit to the data, and the black line is the 1:1 line. Results from the linear regression are given in Table 3.

(backscatter, transmission, etc.). As described in Tucker et al. (2018), for a given atmospheric scattering condition, the GrOAWL instrument performance can be parameterized under instrument throughput efficiency (ITE) and instrument contrast (IC). ITE defines the percentage of total backscattered laser photons collected by the telescope that contribute to the total lidar signal (measurement SNR). IC determines how well the interferometer can resolve Doppler-shifted phases.

The GROAWL ITE is calculated based on efficiency values measured before the flight campaign or provided by the component manufacturers. The GrOAWL IC is the maximum contrast achievable for the interferometer and laser combination. The GrOAWL ITE and IC parameters were first defined for the "as built" instrument, which is unaffected by the operating environment of the instrument (aircraft thermal and vibrational effects). Both parameters were then adjusted to account for platform effects on the instrument performance to simulate the "in-flight" instrument. The in-flight model result was then compared with the observations from the flights and the as-built instrument model result to characterize the impact on instrument performance during the flights.

The total wind measurement performance (precision) depends ultimately on both the instrument performance and on atmospheric parameters including the aerosol scattering ratio and extinction. The large OPD of the GrOAWL interferometer (90 cm) is designed to measure winds from only aerosol backscattered photons (Tucker et al. 2018). As a result, if there are no aerosol-backscattered photons, then the wind measurement contrast and thus the precision would be very low even if both the lidar SNR and the instrument contrast were high. This is akin to longer-wavelength Doppler lidar systems that will see little or no backscattered signal if there are insufficient aerosols.

In the model, the atmospheric temperature and pressure profiles, used to determine molecular backscatter coefficient profiles, were measured by the dropsondes during the flight, and the aerosol extinction coefficient profiles were obtained from the GEOS-5 forecast. GEOS-5 aerosol extinction coefficients were scaled by a scaling factor determined by comparing GEOS-5 aerosol optical depth (AOD) over Houston with AERONET (Holben et al. 2001) AOD measurements in Houston. The GEOS-5 AOD over Houston agreed with the AERONET measurements within 25% on average during the measurement period in June 2016. An aerosol backscatter coefficient profile was then calculated from the total aerosol extinction coefficient using GEOS-5 species dependent lidar ratio given in Nowottnick et al. (2015). GEOS-5 provides aerosol extinction coefficients for five different aerosol species: black carbon, sulfate, sea salt, organic carbon, and dust. The calculated aerosol backscatter profiles were found to be in good agreement with the mean CALIOP backscatter profile over the Gulf of Mexico (26°–29°N, 94°–97°W) for June 2016. The calculated aerosol backscatter profile for the three validation flights together with the mean CALIOP backscatter profile over the Gulf of Mexico are shown in Fig. 9. Because of scheduling conflicts, flights were not performed during CALIOP overpasses.

To assess the instrument throughput performance, the total lidar backscatter signal (digitizer counts translated

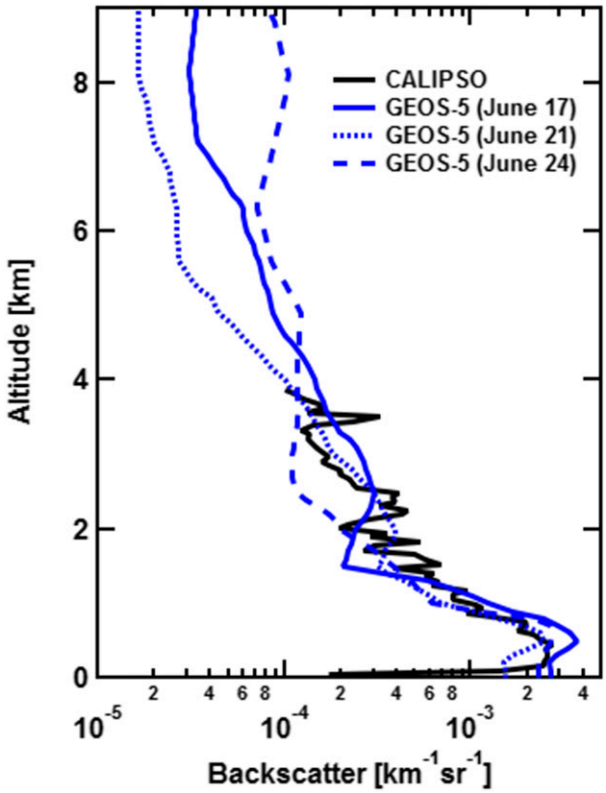

FIG. 9. Aerosol backscatter profiles: derived from mean GEOS-5 1800 UTC forecast extinction profiles along the flight track for the flights on 17 (solid blue), 21 (dotted blue), and 24 Jun 2016 (dashed blue), and mean CALIOP profile over the Gulf of Mexico for June 2016 (black). GEOS-5 extinction coefficients were scaled by AOD comparison with AERONET measurements at Houston. Species-dependent lidar ratio given in Nowottnick et al. (2015) was used to calculate GEOS-5 backscatter profiles.

into average photons per sample) measured during a flight was compared with the simulated total detected lidar backscattered photons based on ideal system performance for the atmospheric conditions encountered. In GrOAWL, the total number of detected lidar photons consists of both the molecular and aerosol backscattered photons from the atmosphere. Because the molecular backscatter coefficient can be more than an order of magnitude larger than the aerosol backscatter coefficient, the total lidar signal is usually dominated by the molecular signal over most of the atmosphere. Thus, for a known transmitted pulse energy the GrOAWL ITE during the flights was determined by adjusting the ITE parameter in the model so that the simulated total lidar backscattered photons match the measurements. Only lidar signals from 1.5–6-km altitude (~3 km below the flight altitude) were evaluated in order to eliminate the impact of the telescope overlap function close to the instrument and higher aerosol loading inside the boundary layer.

Figure 10a shows the 50th percentile of the total measured backscattered (black line) photons as a function of altitude for look 2 during the 17 June 2016 flight. The shaded area covers the 10th–90th percentiles of the data. The variability in the measured photons is due to combination of atmospheric variability (mainly aerosol and cloud extinction and backscatter) and changes in instrument performance (e.g., telescope overlap) over the course of the flight. The corresponding simulation for the in-flight instrument is shown in blue and indicates an average 40% throughput loss on look 2 compared to the as-built instrument (dashed blue). Instrument throughput loss for look 1 during the same flight averaged about 60%. This difference in instrument throughput between the two looks is due to their separate lasers, transmit and receive path alignments, telescope overlap functions, and optical paths to the interferometer. Throughput loss includes both instrumental and environmental effects, such as foggy windows, that would result in lower ITE. Any reduction in laser pulse energy (1.5 mJ) during the flights is also associated with ITE in this analysis.

The platform effects on the average IC were determined by comparing the flight-measured T0 contrast values with those from ground-based testing, where the average GrOAWL T0 contrast was around 0.73, close to the maximum contrast achievable for the interferometer and laser combination (Tucker et al. 2018). The GrOAWL instrument contrast during the flights, however, was reduced because of vibration-induced interferometer misalignment. Average T0 contrasts during the 17 June 2016 flight were 0.47 and 0.33 for look 1 and look 2, respectively. While the same interferometer is used for both looks, the average T0 contrast values for the two looks differ based on differences in the laser pulse bandwidths and T0 signal amplitudes. If T0 signals become too weak (e.g., because of misalignment of the captured T0 signal), then the T0 contrast will also be compromised, as was the case during the flights.

A simulation accounting for losses in both the ITE and IC showed that the instrument performance in terms of median LOSV precision estimates during the AOVT flights (in flight) was reduced by about a factor of 3.3 compared to the as-built (ground test) instrument, where the instrument contrast was 0.73. Figure 10b shows the CNR-based precision estimates [Eq. (3)] for the actual in-flight GrOAWL measurements (black line and gray shading), the in-flight instrument model accounting for platform-induced losses (solid blue line), and the as-built instrument model (dashed blue line). The impacts of ITE (i.e., fewer photons collected) and IC losses (i.e., interferometer alignment) on instrument precision were similar in magnitude. Performance was similar for all

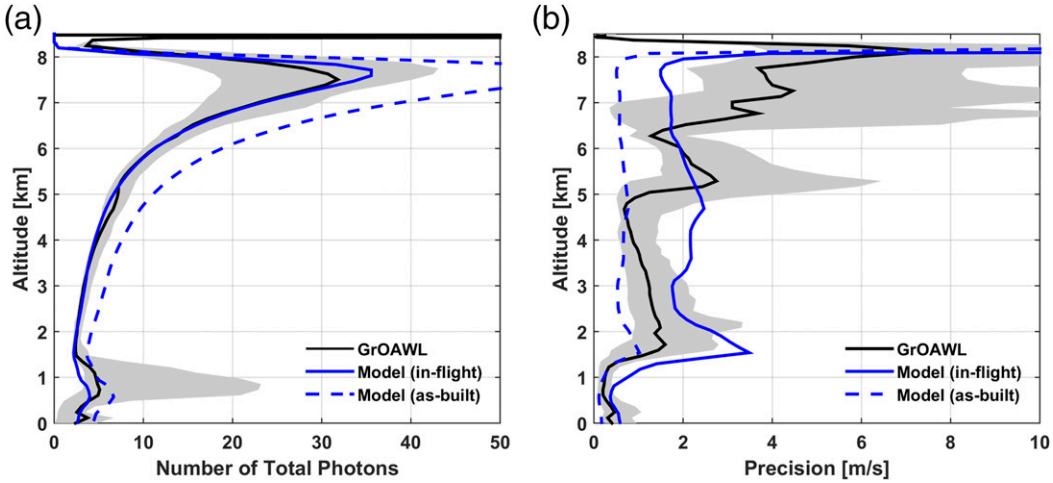

FIG. 10. (a) Total signal photons (Rayleigh and aerosol): The median of the total signal photons measured by the GrOAWL look 2 during the flight on 17 Jun 2016 (black). Gray shading represents the 10th and 90th percentiles of the data. Simulated total photons for the (i) as-built instrument (dashed blue) and (ii) in-flight instrument, which account for 40% throughput reduction (solid blue). (b) LOSV precision: median LOSV precision for the same flight (black). The gray shading bounds the 10th- and 90th-percentile data ranges. The simulated instrument precisions are shown for (i) as-built instrument (dashed blue) and (ii) in-flight instrument (solid blue). Temperature and pressure profiles used in the simulation were measured by a dropsonde during the same flight, and the aerosol backscatter coefficient profile (see Fig. 9) was derived from a GEOS-5 extinction coefficients profile scaled by AERONET AOD measurement. Surface return is not simulated in the model.

three validation flights. The agreement between the measured and in-flight model precision profiles does not show as good of an agreement as the total number of photons shown in Fig. 10a. This is due to differences in the GEOS-5-derived aerosol backscatter profile and the true measurement backscatter environment, which includes both aerosols and clouds. However, since evaluation of instrument performance looks at ITE (using information about the molecular return signal amplitudes) and measured IC (T0 contrast), and compares the precision estimates from as-built and in-flight instrument models for the same aerosol backscatter coefficient profile, the instrument performance model analysis results are independent of the aerosol backscatter profile used in the model.

After the AOVT flights, the GrOAWL instrument was set up in a Ball Aerospace rooftop laboratory in Boulder, Colorado, where it was shown to perform close to the predicted ideal performance for known aerosol conditions and as-built instrument parameters, indicating that aircraft operating environment did degrade the instrument measurement performance during the flights. Vibration and temperature variations in the laboratory setting were minimal compared to the WB-57 research flights. Aerosol information was provided by the NCAR GV HSRL instrument (UCAR/NCAR 2017). Results from the ground-based laboratory measurements will be presented in a separate publication. The engineering solutions to the in-flight issues being incorporated in the next-generation OAWL instrument should result in much improved in-flight airborne performance in the future. More importantly for deployment on space-based platforms, issues such as temperature variations and vibration levels are smaller, more consistent, and better understood, and have already been mitigated for spaceborne lidar instruments.

5. Implications for space-based wind measurements

The GrOAWL instrument was originally built to demonstrate and validate the ATHENA-OAWL satellite mission concept, but the results from the AOVT flight tests are not limited to scaling to the ATHENA-OAWL mission parameters. The results can be scaled to different OAWL-based mission architectures, based on mission requirements and available platforms. The main differences between a space-based OAWL (SpOAWL) instrument and the airborne GrOAWL instrument are the laser power P (pulse energy times pulse repetition frequency), telescope diameter d, range to the measurement region of interest R, and processing parameters. To compensate for (scale up to) the larger ranges to the measurement region of interest from space, SpOAWLs use higher-powered lasers, larger telescopes, and longer accumulation times and range gates for data

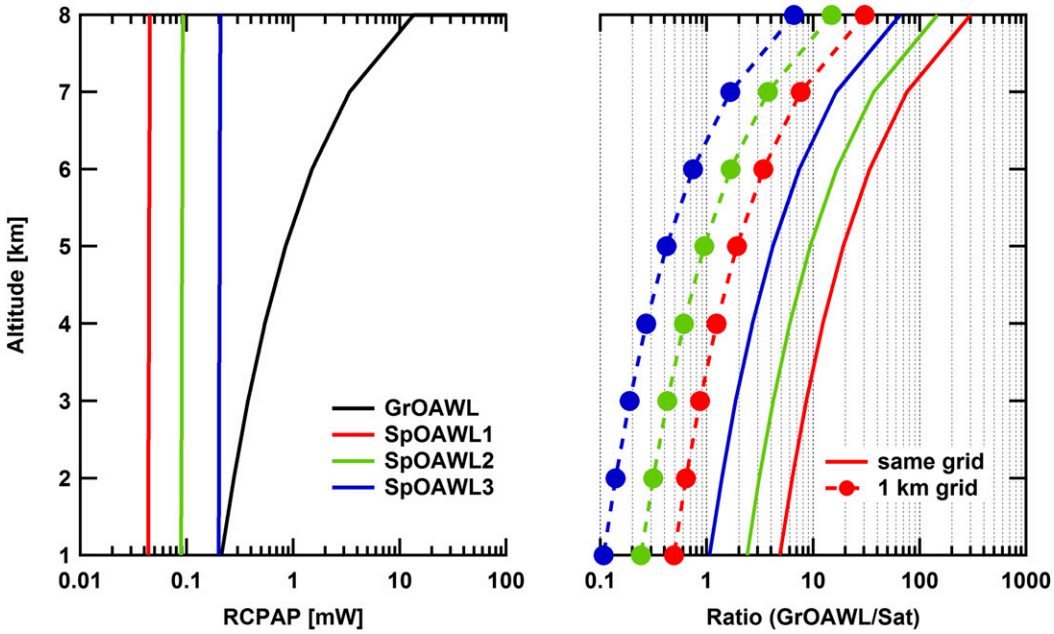

FIG. 11. (a) RCPAP for the as-built GrOAWL system during the AOVT flight campaign (black) and three different SpOAWL instruments (other colors). Space-based instruments differ only in the telescope size. SpOAWL1 configuration is identical to the proposed EVI-2 ATHENA-OAWL mission and includes a 0.7-m-diameter telescope. SpOAWL2 and SpOAWL3 use 1.0- and 1.5-m-diameter telescopes, respectively. (b) The ratio of RCPAP between the GrOAWL and SpOAWL instruments for the same 121-m vertical grid and 10-s time resolution retrieval (solid lines), and 1-km vertical grid and 12-s time resolution retrieval for the SpOAWL instruments (dashed lines with closed circles).

processing. The range-corrected power aperture product (RCPAP; Pd^2R^{-2}), for three different SpOAWL configurations and the GrOAWL system, demonstrating the instrument differences over the demonstrated aircraft flight altitude ranges, is shown in Fig. 11a. Table 1 lists instrument parameters for the GrOAWL and the three SpOAWL instruments. One of the SpOAWL instrument configurations (SpOAWL1) represents the proposed ATHENA-OAWL mission for EVI-2. The other two configurations differ from SpOAWL1 only in telescope size. The sizes are representative of the telescopes currently being used in other space-based lidars [e.g., 1.0 m for CALIOP (Hunt et al. 2009), and 1.5 m for Aeolus (Stoffelen et al. 2005)]. Because of the shorter ranges from the aircraft platform, the GrOAWL RCPAP shows more of the inverse-range-squared dependence compared to the near-constant values for the satellite instruments over the altitudes of interest, a result of the long range from orbit to the lower troposphere. For measurements from 4-km altitude, depending on telescope size, RCPAP values are about 3–12 times larger for the GrOAWL instrument than for the satellite instruments (Fig. 11b, solid colored lines). Lidar return signal strength scales linearly with laser and receiver efficiency (including detector quantum efficiency) and quadratically with telescope diameter. While telescope size is shown here as an example, other instrument parameters, such as laser pulse energy, can also be adjusted to scale an instrument's performance to meet the requirements of a space-based mission.

Although the RCPAP values are lower because of the long ranges from orbit, the space-based instruments' measurement performance requirements can be met by accumulating data both vertically and horizontally. This larger binning for space [1000-m vertical grid and 12 s (1800 pulses, ~90 km) horizontal] compared to the nominal airborne GrOAWL binning [~120 m and 10 s (2000 pulses)] compensates for the lower space-based instruments' RCPAP values in the lowest 4 km, enabling them to be more precise than the GrOAWL at those altitudes (Fig. 11b, dashed colored lines with closed circles). Thus, if any of the SpOAWL instruments were making measurements at the same location and time as the AOVT flights, then the precision of the measured winds at 1-km vertical grid would be comparable or better than the GrOAWL winds' precision at altitudes below 4 km; that is, the precision of the SpOAWL-measured winds on a 1-km vertical and ~90-km horizontal grid would be comparable or better than the GrOAWL winds' precision at altitudes below 4 km.

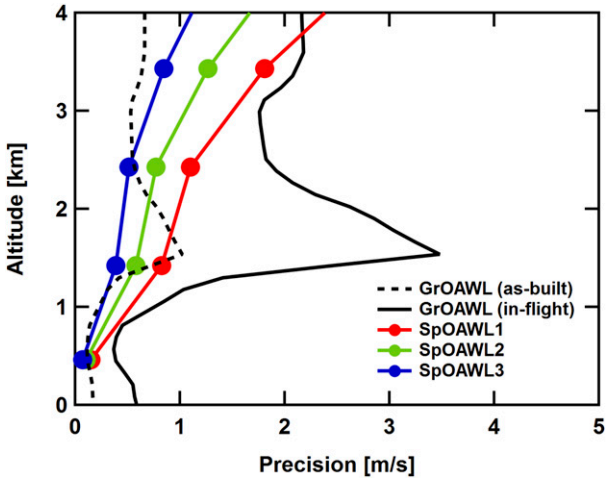

FIG. 12. Simulated instrument precision for the GrOAWL (black) and three SpOAWL instrument architectures (color) using the GEOS-5 aerosol backscatter profile from 17 Jun 2016 (see Fig. 9, solid blue line). The dashed black line represents the as-built GrOAWL instrument, and the solid black line shows the in-flight GrOAWL instrument (as in Fig. 10, right panel, solid blue line). Precision values are calculated for ~121-m vertical grid and 10-s (2000 pulse) time accumulation for the GrOAWL, and 1000-m vertical grid and 12-s (1800 pulse) time accumulation for the SpOAWL instruments.

A comparison of system performance above 4 km is difficult to infer from the GrOAWL data because of the nonlinearity of the inverse-range-squared correction and telescope overlap function; the GrOAWL RCPAP values are much larger closer to the aircraft-mounted instrument.

Figure 12 shows the expected precision estimate for the three SpOAWL instruments for the atmospheric conditions observed during the 17 June flight. The instrument performance model, validated using the AOVT flights results, was used to simulate the space-based instruments' performance for the same conditions. The SpOAWLs' precision (colored lines) are very similar to the as-built GrOAWL instrument precision (dotted black line) up to 4-km altitude for the same GEOS-5 aerosol scenario (see Fig. 9). The GrOAWL performance is relatively better than the SpOAWLs above 4 km (not shown) because of its larger RCPAP values closer to the airborne instrument. The in-flight GrOAWL instrument precision (solid black line) indicates potential leeway in the SpOAWL instrument's parameters while still being able to make good measurements. The actual instrument performance from space would depend upon the specific instrument's configuration and aerosol backscatter conditions (a highly variable parameter) during the time of the measurement. Thus, an OAWL system with lower laser pulse energy (<160 mJ used for SpOAWLs) would still be able to make precise measurements from space in areas of higher aerosol loading. Sampling error caused by spatial variability in aerosol backscatter and wind over the large measurement volume for space-based measurements is likely to be larger than instrument-induced error. For any design point, however, wind speed measurements with a system designed for aerosol returns will continue to be challenging under extremely low-to-no-aerosol conditions, suggesting the need for a complementary molecular Doppler wind lidar channel (e.g., Lux et al. 2018; Stoffelen et al. 2005) to obtain profiles throughout the troposphere by measuring winds from these low-backscatter regions. The OAWL measurement concept enables the addition of a complementary molecular channel to the existing aerosol-based instrument to simultaneously measure winds from aerosol and molecular backscatter using a single system. Current efforts are underway to develop and demonstrate the molecular channel for OAWL so that a future space-based system could measure winds throughout the troposphere, from aerosol and molecular returns.

6. Summary and conclusions

The ATHENA-OAWL airborne demonstrator instrument, GrOAWL, was deployed aboard the NASA WB-57 research aircraft along with an automated dropsonde system during the AOVT flight campaign. The goal of the campaign was to demonstrate the instrument performance and the wind measuring capability. Harsh operating conditions of the WB-57 aircraft caused system underperformance during the flights. Simulations of the GrOAWL performance indicated about a factor of 3.3 reduction in instrument LOSV measurement precision (3.3-times increase in uncertainty) during the flights, relative to expected ideal performance. Despite these issues, the instrument operated autonomously and accurately measured LOSVs using both looks, concurrently. The GrOAWL-measured LOSVs for both looks, as well as wind speed and direction retrieved by combining the two LOSVs, were validated by comparison with the dropsonde measurements. The agreement between the two measurements was excellent ($R^2 > 0.9$) with very small bias (look 1: $-0.07 \pm 0.07\,\mathrm{m\,s^{-1}}$ and look 2: $0.01 \pm 0.07\,\mathrm{m\,s^{-1}}$). This is the first concurrent two-look DWL system for continuous wind profile measurement, and it successfully demonstrated the proposed two-look architecture for a space-based platform. Scaling the GrOAWL performance to space using the GrOAWL instrument performance model showed space-based OAWL instruments are capable of providing wind measurements with precision comparable to and better than those from the GrOAWL system during the AOVT flights.

These results from the flight campaign provide strong evidence that mission concepts like that proposed for ATHENA-OAWL are suitable for measurement of winds from space. Hardware problems encountered during the flights were specific to the operating environment of the aircraft and should not be an issue for the spaceborne measurements. Engineering solutions to these hardware issues are already being implemented to the updated version of the instrument (HAWC-OAWL). The potential to make HSRL measurements of aerosol and clouds concurrently with winds with the single-OAWL instrument makes the OAWL concept even more attractive for spaceborne measurements.

Acknowledgments. The authors acknowledge the excellent efforts of the Ball Aerospace OAWL team, especially Rachel Narcisco, Mike Adkins, Mike Conde, Jeff Applegate, Noah Siegel, Glen Taudien, Paul Kaptchen, and Carl Weimer. The authors thank the NASA flight and support crew at Ellington Field for all their support during the flight campaign, the NWS Corpus Christi office for launching additional radiosonde during the flights, Dave Emmitt for the helpful discussions during planning and data analysis, and James Flynn for establishing and maintaining the University of Houston AERONET site. The GEOS-5 data have been provided by the Global Modeling and Assimilation Office at NASA Goddard Space Flight Center through the online data portal in the NASA Center for Climate Simulation. The CALIOP data were obtained from the NASA Langley Research Center Atmospheric Science Data Center. This work was funded by the NASA ESTO and NASA ESSP Venture Technology program (NNX15AE57G, NNX15AE59G). Additional support for the dropsondes was provided by the Office of Naval Research.

REFERENCES

Baidar, S., and Coauthors, 2013: Combining active and passive airborne remote sensing to quantify NO2 and Ox production near Bakersfield, CA. *Br. J. Environ. Climate Change*, **3**, 566–586, https://doi.org/10.9734/bjecc/2013/5740.

Baker, W. E., and Coauthors, 2014: Lidar-measured wind profiles: The missing link in the global observing system. *Bull. Amer. Meteor. Soc.*, **95**, 543–564, https://doi.org/10.1175/BAMS-D-12-00164.1.

Bilbro, J., G. Fichtl, D. Fitzjarrald, M. Krause, and R. Lee, 1984: Airborne Doppler lidar wind field measurements. *Bull. Amer. Meteor. Soc.*, **65**, 348–359, https://doi.org/10.1175/1520-0477(1984)065<0348:ADLWFM>2.0.CO;2.

Black, P., L. Harrison, M. Beaubien, R. Bluth, R. Woods, A. Penny, R. W. Smith, and J. D. Doyle, 2017: High-Definition Sounding System (HDSS) for atmospheric profiling. *J. Atmos. Ocean. Technol.*, **34**, 777–796, https://doi.org/10.1175/JTECH-D-14-00210.1.

Browning, K. A., and R. Wexler, 1968: The determination of kinematic properties of a wind field using Doppler radar. *J. Appl. Meteor.*, **7**, 105–113, https://doi.org/10.1175/1520-0450(1968)007<0105:TDOKPO>2.0.CO;2.

Bruneau, D., and J. Pelon, 2003: Simultaneous measurements of particle backscattering and extinction coefficients and wind velocity by lidar with a Mach–Zehnder interferometer: Principle of operation and performance assessment. *Appl. Opt.*, **42**, 1101–1114, https://doi.org/10.1364/AO.42.001101.

Chouza, F., O. Reitebuch, A. Benedetti, and B. Weinzierl, 2016: Saharan dust long-range transport across the Atlantic studied by an airborne Doppler wind lidar and the MACC model. *Atmos. Chem. Phys.*, **16**, 11 581–11 600, https://doi.org/10.5194/acp-16-11581-2016.

Gisi, M., F. Hase, S. Dohe, T. Blumenstock, A. Simon, and A. Keens, 2012: XCO_2-measurements with a tabletop FTS using solar absorption spectroscopy. *Atmos. Meas. Tech.*, **5**, 2969–2980, https://doi.org/10.5194/amt-5-2969-2012.

Grund, C. J., J. Howell, R. Pierce, and M. Stephens, 2009: Optical autocovariance direct detection lidar for simultaneous wind, aerosol, and chemistry profiling from ground, air, and space platforms. *Proc. SPIE*, **7312**, 73120U, https://doi.org/10.1117/12.824204.

Holben, B. N., and Coauthors, 2001: An emerging ground-based aerosol climatology: Aerosol optical depth from AERONET. *J. Geophys. Res.*, **106**, 12 067–12 097, https://doi.org/10.1029/2001JD900014.

Holton, J. R., 2004: *An Introduction to Dynamic Meteorology*. Academic Press, 535 pp.

Horányi, A., C. Cardinali, M. Rennie, and L. Isaksen, 2015: The assimilation of horizontal line-of-sight wind information into the ECMWF data assimilation and forecasting system. Part II: The impact of degraded wind observations. *Quart. J. Roy. Meteor. Soc.*, **141**, 1233–1243, https://doi.org/10.1002/qj.2551.

Hunt, W. H., D. M. Winker, M. A. Vaughan, K. A. Powell, P. L. Lucker, and C. Weimer, 2009: *CALIPSO* lidar description and performance assessment. *J. Atmos. Oceanic Technol.*, **26**, 1214–1228, https://doi.org/10.1175/2009JTECHA1223.1.

Liu, Z., and T. Kobayashi, 1996: Differential discrimination technique for incoherent Doppler lidar to measure atmospheric wind and backscatter ratio. *Opt. Rev.*, **3**, 47–52, https://doi.org/10.1007/s10043-996-0047-0.

Lux, O., C. Lemmerz, F. Weiler, U. Marksteiner, B. Witschas, S. Rahm, A. Schäfler, and O. Reitebuch, 2018: Airborne wind lidar observations over the North Atlantic in 2016 for the prelaunch validation of the satellite mission Aeolus. *Atmos. Meas. Tech.*, **11**, 3297–3322, https://doi.org/10.5194/amt-11-3297-2018.

NASEM, 2018: *Thriving on Our Changing Planet: A Decadal Strategy for Earth Observation from Space*. National Academies Press, 700 pp., https://doi.org/10.17226/24938.

Nowottnick, E. P., P. R. Colarco, E. J. Welton, and A. Da Silva, 2015: Use of the CALIOP vertical feature mask for evaluating global aerosol models. *Atmos. Meas. Tech.*, **8**, 3647–3669, https://doi.org/10.5194/amt-8-3647-2015.

Paffrath, U., C. Lemmerz, O. Reitebuch, B. Witschas, I. Nikolaus, and V. Freudenthaler, 2009: The airborne demonstrator for the direct-detection Doppler wind lidar ALADIN on ADM-Aeolus. Part II: Simulations and Rayleigh receiver radiometric performance. *J. Atmos. Oceanic Technol.*, **26**, 2516–2530, https://doi.org/10.1175/2009JTECHA1314.1.

Reitebuch, O., C. Lemmerz, E. Nagel, U. Paffrath, Y. Durand, M. Endemann, F. Fabre, and M. Chaloupy, 2009: The airborne demonstrator for the direct-detection Doppler wind lidar ALADIN on ADM-Aeolus. Part I: Instrument design and comparison to satellite instrument. *J. Atmos. Oceanic Technol.*, **26**, 2501–2515, https://doi.org/10.1175/2009JTECHA1309.1.

Schwiesow, R. L., and S. D. Mayor, 1995: Coherent optical signal processing for a Doppler lidar using a Michelson interferometer. *Coherent Laser Radar*, OSA Technical Digest Series, Vol. 19, Optical Society of America, 212–215.

Stoffelen, A., and Coauthors, 2005: The atmospheric dynamics mission for global wind field measurement. *Bull. Amer. Meteor. Soc.*, **86**, 73–88, https://doi.org/10.1175/BAMS-86-1-73.

Targ, R., M. J. Kavaya, R. M. Huffaker, and R. L. Bowles, 1991: Coherent lidar airborne windshear sensor: Performance evaluation. *Appl. Opt.*, **30**, 2013, https://doi.org/10.1364/AO.30.002013.

Tucker, S. C., C. S. Weimer, and R. M. Hardesty, 2016: The ATHENA-OAWL Doppler wind lidar mission. *EPJ Web Conf.*, **119**, 01002, https://doi.org/10.1051/epjconf/201611901002.

——, ——, S. Baidar, and R. M. Hardesty, 2018: The optical autocovariance wind lidar (OAWL). Part I: Instrument development and demonstration. *J. Atmos. Oceanic Technol.*, **35**, 2079–2097, https://doi.org/10.1175/JTECH-D-18-0024.1.

UCAR/NCAR, 2017: Gulfstream-V High Spectral Resolution Lidar. Accessed 7 September 2017, https://doi.org/10.5065/D67W6976.

Wang, J., J. Bian, W. O. Brown, H. Cole, V. Grubišić, and K. Young, 2009: Vertical air motion from T-REX radiosonde and dropsonde data. *J. Atmos. Oceanic Technol.*, **26**, 928–942, https://doi.org/10.1175/2008JTECHA1240.1.

Weimer, C., S. Tucker, W. Baker, R. M. Hardesty, and L. P. Riishojgaard, 2015: The ATHENA-OAWL Doppler wind lidar mission. *Fifth Conf. on Transition of Research to Operations*, Phoenix, AZ, Amer. Meteor. Soc., J4.4, https://ams.confex.com/ams/95Annual/webprogram/Paper268581.html.

Witschas, B., S. Rahm, A. Dörnbrack, J. Wagner, and M. Rapp, 2017: Airborne wind lidar measurements of vertical and horizontal winds for the investigation of orographically induced gravity waves. *J. Atmos. Oceanic Technol.*, **34**, 1371–1386, https://doi.org/10.1175/JTECH-D-17-0021.1.

WMO, 2016a: Statement of guidance for global numerical weather prediction (NWP). Accessed 31 March 2017, 12 pp., http://www.wmo.int/pages/prog/www/OSY/SOG/SoG-Global-NWP.pdf.

——, 2016b: Statement of guidance for high-resolution numerical weather prediction (NWP). Accessed 31 March 2017, 10 pp., http://www.wmo.int/pages/prog/www/OSY/SOG/SoG-HighRes-NWP.pdf.

Zhang, J., R. Atlas, G. Emmitt, L. Bucci, and K. Ryan, 2018: Airborne Doppler wind lidar observations of the tropical cyclone boundary layer. *Remote Sens.*, **10**, 825, https://doi.org/10.3390/rs10060825.