Impact of airborne Doppler lidar observations on ECMWF forecasts

Martin Weissmann[a]* and Carla Cardinali[b]

[a] *Institut für Physik der Atmosphäre, DLR Oberpfaffenhofen, Germany*
[b] *European Centre for Medium-Range Weather Forecasts, Reading, UK*

ABSTRACT: During the Atlantic THORPEX Regional Campaign (A-TReC) in autumn 2003, the airborne Doppler lidar of the Deutsches Zentrum für Luft- und Raumfahrt (DLR) was used to observe wind in predicted sensitive regions. In eight flights, the system measured a total of 1600 wind profiles that were experimentally assimilated in the global assimilation system at the European Centre for Medium-Range Weather Forecasts (ECMWF). The present study assesses the impact of these Doppler lidar measurements in Numerical Weather Prediction (NWP). It is shown that lidar observations have a significant impact on the analyses as well as on forecasts owing to high accuracy and spatial resolution. The measurements reduce the errors of the 1–4 day forecasts of geopotential height, wind, and humidity over Europe throughout the troposphere. On average, Doppler lidar measurements reduce the 2–4 day forecast error of geopotential height over Europe by 3%. This is a promising result, considering that observations have been gathered from only 28.5 flight hours. Dropsondes released in the same area as the Doppler lidar was operating show good agreement for measured winds, but smaller analysis impact and less reduction of the forecast error. Copyright © 2007 Royal Meteorological Society

1. Introduction

The wind field over oceans is still poorly observed. Single-level wind measurements are provided at the surface by buoys, ships, and satellite scatterometers, while aircraft observe wind mainly at the cruise level along the air traffic corridors. Wind profiles are only provided by a small number of radiosondes launched from ships. Satellite cloud-drift winds are numerous, but they have fairly large errors owing to inaccurate height assignment. In such a framework, Doppler wind lidar offers a great opportunity to measure the wind field either globally, with a polar-orbiting satellite (Tan and Andersson, 2005), or regionally, if mounted on aircraft.

As part of THORPEX (a Global Atmospheric Research Programme [http://www.wmo.int/thorpex, accessed on 24 November 2006]), additional measurements were taken over the North Atlantic between mid-October and mid-December 2003 to complement the operational global observing system. The main objective of this Atlantic THORPEX Regional Campaign (A-TReC) was to reduce the 1–3 day forecast error of Numerical Weather Prediction (NWP) models over Europe and northeastern America by better observing the initial state of the atmosphere (Rabier *et al.*, 1996). The additional measurements include dropsondes launched from research

aircraft, AMDAR (Aircraft Meteorological DAta Relay), ASAP (Automated Shipboard Aerological Programme), radiosondes, drifting buoys, satellite rapid scan winds, and the airborne wind lidar of the Deutsches Zentrum für Luft- und Raumfahrt (DLR). In a joint effort operational NWP centres, universities, and research institutes evaluated the impact of the observations. These studies will provide further insights into the value of different observation systems and a basis for the optimization of the global observation system. Results on the impact of A-TReC data (except lidar) were provided for large verification areas (Langland, 2005) and for the specific verification regions (Rabier, 2006).

This study assesses the impact of airborne Doppler lidar measurements on the analyses and forecasts of the European Centre for Medium-Range Weather Forecasts (ECMWF), and compares their impact with the impact of 97 dropsondes deployed in the same time period. The airborne Doppler lidar operated for a 2 week period and over eight flights; a total of ~1600 wind profiles was recorded. About half of the measurements were taken in sensitive areas with the aim of improving the forecast of a specific weather event in a regional verification area. However, the present impact study focuses on the impact over Europe during the 2 weeks. The investigation of the individual targeted cases will be the subject of subsequent studies.

The lidar measurements have been assimilated in the global assimilation system at ECMWF. Several

* Correspondence to: Martin Weissmann, Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Physik der Atmosphäre, D-82230 Wessling, Germany. E-mail: martin.weissmann@dlr.de

experiments, with different averaging methods applied to the lidar data and different observation error, were carried out to investigate how such measurements should best be assimilated. The lidar measurements were averaged along the flight track, making them far more representative of the wind field in a grid box than, for example, dropsondes and rawinsondes. The instrumental error of airborne Doppler lidars is significantly smaller than could be achieved with a satellite instrument. In consequence, the total lidar observation error (representativeness and instrumental errors together) is smaller than the error of all other wind measurements in the current operational observation system. More information about the airborne lidar measurements during A-TReC and the lidar system can be found online at http://www.pa.op.dlr.de/na-trec (accessed on 24 November 2006).

The outline of this paper is as follows: section 2 describes the lidar system, the ECMWF global model, and the assimilation procedure; section 3 discusses the experiments performed; and section 4 summarizes the results and gives an outlook of further research in this area and possible applications of airborne wind lidars.

2. Description of measurements and the model system

2.1. The scanning airborne Doppler lidar system of DLR

The principle of Doppler lidar is in many ways similar to that of Doppler radars except that a lidar emits pulses of laser light instead of radio waves (Grund *et al.*, 2001). The airborne DLR Doppler lidar system measures wind profiles beneath the aircraft using the Velocity-Azimuth Display (VAD) technique. The instrument performs a conical step-and-stare scan around the vertical axis at 20° off nadir. At 24 different azimuth angles, the lidar emits 500 or 1000 laser pulses at a wavelength of 2.02254 µm and measures the backscatter signal from atmospheric aerosol. The Doppler shift of the backscatter signal is proportional to the velocity component in the pointing direction, the line-of-sight (LOS) velocity. The distance of the measurement volume from the aircraft is a function of the time difference between emitting the laser pulse

and receiving the backscattered signal. Consequently, profiles of the LOS velocity are measured at 24 different azimuth angles per scanner revolution.

The wind vectors are derived from the LOS measurements usually at 24 or 96 different azimuth angles using a combination of an inversion and a maximum-likelihood algorithm (Smalikho, 2003; Weissmann *et al.*, 2005b).

Combined with the movement of the aircraft, the conical scanning leads to a cycloid scan pattern beneath the aircraft and the derived winds are averaged along this scan. This horizontal averaging is the main advantage of airborne Doppler lidar observations compared with *in situ* observations as it makes the data more representative of the wind in a model grid box. The horizontal width of the scan pattern increases with the distance from the aircraft: at 5 and 10 km beneath the aircraft the width is about 3.5 and 7 km, respectively. During A-TReC, the horizontal length of one scanner revolution was usually ∼10 km. Thus the wind profiles derived from 24 LOS velocity profiles have a horizontal resolution of ∼10 km, and the profiles derived from 96 LOS velocity profiles have a resolution of ∼40 km. The vertical resolution of the wind profiles is 100 m.

The hardware of the lidar system consists of a transceiver developed by Lockheed Martin Coherent Technologies (http://www.lockheedmartin.com/coherent/), a rotating silicon wedge scanner, and a data acquisition system developed at DLR (Köpp *et al.*, 2004).

The same airborne lidar system has already been used in the German field campaign VERTIKATOR (Weissmann *et al.*, 2005a), and another one has been used to study the flow over the Alps (Reitebuch *et al.*, 2003), the French Mistral (Drobinski *et al.*, 2004), and a land–sea breeze in southern France (Bastin *et al.*, 2005).

2.2. Lidar measurements during A-TReC

The airborne Doppler lidar operated from the DLR Falcon aircraft during eight flights between 14 and 28 November 2003. Figure 1 shows the flight tracks obtained. Four of these flights were part of targeted campaigns and directed into sensitive areas, two were transfer flights (14 and 28 November), one was directed towards Greenland to
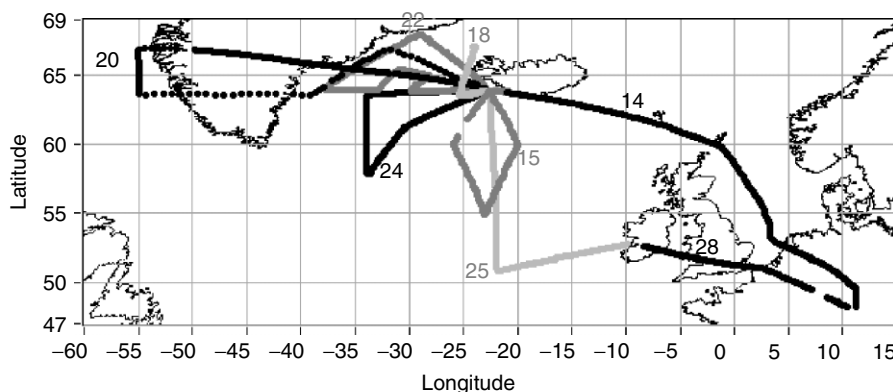

Figure 1. Flight tracks of all flights during A-TReC. Numbers indicate the date of the flights in November 2003.

measure the lee effects of the island (24 November), and one was designed for an intercomparison of lidar winds and the Advanced Synthetic Aperture Radar (ASAR) on the European Environmental Satellite ENVISAT (18 November).

The data were processed with a horizontal resolution of ~10 km (one scanner revolution) and in parallel also with ~40 km (four scanner revolutions) to increase the measurement coverage and to reduce representativeness errors. The 10 km dataset consists of ~1600 vertical profiles with a total of ~40 000 individual wind measurements. This means that there are on average 25 reliable wind measurements on every profile. The 40 km dataset contains ~400 profiles with a total of ~15 000 wind measurements. This is equivalent to an average of 38 reliable wind measurements per profile.

The flight altitude varied between 2800 and 11 000 m mean sea level (MSL). The lidar measurements start at a distance of 500 m from the aircraft, which leads to lidar measurements ranging from 0 to 10 500 m MSL. There was a large variation of the lidar coverage at different heights in the individual campaigns, but the total amount of measurements during the 2 weeks was fairly evenly distributed over all vertical levels.

The vertical cross section along the aircraft flight track counts 36% and 54% data coverage for the 10 km and 40 km resolution, respectively. Missing values in the data are either due to clear air with low aerosol content or to optically thick clouds that cannot be penetrated by the lidar.

The standard deviation (STD) of the instrumental error was determined to be in the range of 0.75–1 m s^{-1} through an intercomparison of collocated dropsonde and lidar measurements. The total error for the assimilation of such measurements (instrumental and representativeness errors) was estimated to be 1–1.5 m s^{-1} for all model levels. This is smaller than the errors of all current routine observations: at ECMWF the STD error of 1.8–3 m s^{-1} is assigned to radiosonde and dropsonde wind observations, 2.5–3.4 m s^{-1} to aircraft measurements, and 2–5.7 m s^{-1} to satellite cloud-drift winds (see http://www.ecmwf.int/research/ifsdocs/CY28r1/Observations, accessed on 24 November 2006)

A detailed description of the intercomparison of lidar and dropsonde measurements together with a comprehensive overview of the lidar measurements during A-TReC, and a full explanation of the airborne Doppler lidar system can be found in Weissmann *et al.* (2005b).

2.3. Model system

The experiments were performed using the ECMWF Integrated Forecast System (IFS) 12 h window 4D-Var. The full resolution (T511, ~40 km) is used for comparing observations and model equivalents and T95/T159 resolution (incremental approach, Courtier *et al.*, 1994; Veersé and Thépaut, 1998) to minimise the departures (Rabier *et al.*, 2000).

One of the strengths of 4D-Var is its ability to assimilate frequent observations (Andersson *et al.*, 2001). All measurements within a 12 h period are used simultaneously in one global estimation problem. The observations are compared every half hour with a short-range forecast. The differences between observations and the short-term forecast (background departures) are minimized to obtain the best unbiased estimate of the initial condition.

At the beginning of the assimilation procedure the system performs a first-guess check (Järvinen and Undén, 1997; Andersson *et al.*, 2000) to reject likely wrong observations, whose background departures exceed a multiple of its expectation: $\exp = \alpha \sqrt{(\sigma_b^2 + \sigma_o^2)}$. Here σ_o and σ_b are the estimated STD of the observation and background errors, respectively. The value of α is set to 5 for most data types, including aircraft and lidar observations.

The measurements are thinned to avoid potential imbalances (Cardinali *et al.*, 2003). After the first-guess check, a variational quality control (VarQC) is performed in the minimization procedure (Andersson and Järvinen, 1999) to decrease the weight of remaining doubtful measurements.

The background and analysis departures of every observation are stored for subsequent diagnostics. Measurements discarded by the first-guess check, by the thinning, or by VarQC are flagged and do not influence the analysis further. Section 3.2 discusses the number of discarded lidar measurements and their background and analysis departures.

3. Experiments

3.1. Experiment definitions

In this study, six experiments were conducted (Table I). A control run (Control) was performed for the whole period of A-TReC with all routine observations. All other experiments started from Control and used either lidar or dropsonde observations in addition to the routine observations. In all lidar experiments the dropsondes were used passively, which means that their background and analysis departures are computed but the measurements are not assimilated. In the experiment 1Rev, lidar wind profiles at a horizontal resolution of ~10 km (the standard resolution of the lidar measurements) were assimilated. Two experiments were performed by assimilating lidar observations horizontally averaged to ~40 km to match the model resolution: the Median experiment assimilated the median of four wind profiles, each one calculated from 24 LOS measurements (one scanner revolution), while 4Rev used wind profiles derived by fitting 96 LOS measurements obtained from four scanner revolutions. The purpose of the Median and 4Rev was to investigate the increase in the representativeness of the measurements obtained by averaging the data horizontally to the model resolution.

In all these experiments an error STD of 1 m s^{-1} was assigned to the lidar data at all vertical levels. This error

Table I. Overview of experiments.

Experiment	Time period	Used measurements and characteristics of experiments
Control	17 October – 15 December	All measurements of the operational analysis without special A-TReC data (dropsondes, AMDAR, etc.)
1Rev	14–30 November	Lidar wind from 1 scan revolution, horiz. resolution $\sim$10 km, assigned STD = 1 ms^{-1}
Median	14–30 November	Median of lidar winds from 1 scan revolution, horiz. res. $\sim$40 km, assigned STD = 1 ms^{-1}
4Rev	14–30 November	Lidar wind from 4 scan revolutions, horiz. resolution $\sim$40 km, assigned STD = 1 ms^{-1}
4RStd	14–30 November	Same lidar measurements as in 4Rev, but higher measurement error: assigned STD = 1.5 ms^{-1}
Drops	14–30 November	Dropsonde wind and temperature measurements, 97 sondes, assigned STD (wind) = 2–3 m s^{-1}

was derived by an intercomparison of collocated lidar and dropsonde measurements, which showed that there is no significant correlation of accuracy and height (Weissmann *et al*., 2005b).

The derived error assumes a continuous lidar measurement through a grid box, whereas in reality, there were gaps in the measurements due to low aerosol concentrations or clouds. Therefore, the experiment 4RStd was performed using the same pre-processed lidar data as 4Rev, but with an error of 1.5 m s^{-1} to take gaps in the observations into account.

The experiment Drops assimilated 97 dropsondes (wind and temperature profiles) released on 10 flights in the same time period. About half of the dropsondes were launched from the DLR Falcon and coincided with lidar measurements. The other half were launched by other research aircraft. The flight time spent for the dropsonde observations was similar to the flight time for the lidar measurements. Thus the cost of the two observation types

is roughly comparable. All these experiments were run from 14 November until 30 November 2003.

3.2. Assimilation statistics

To investigate the performance of Doppler lidar measurements in the analysis, their background departures from the eight analyses with lidar measurements are sampled and compared with the background departures of aircraft and dropsonde measurements in the same area and at the same time. Figure 2 shows the STD of the background departures, the percentage of measurements discarded by the first-guess check, and VarQC for the lidar data in the experiments 1Rev, Median, 4Rev, and 4RStd, respectively, while the dropsonde and aircraft statistics are from 4RStd. For all experiments the STD of the background departures of all lidar measurements is $\sim$3 m s^{-1}, going down to nearly 2.5 m s^{-1} when only used observations are considered, but not the ones discarded by the first-guess check or VarQC. The variability of lidar background departures is smaller than that of aircraft and dropsonde measurements in the same region. This confirms that lidar winds are more representative of the wind per model grid box.

The vertical profile of the STD of lidar background departures (Figure 3) shows the largest values between 850 and 550 hPa. This might be due to a weak backscatter signal from these levels. Usually, the lidar receives a strong signal from the air in the vicinity of the aircraft as well as from the ABL with higher concentrations of aerosols and humidity. However, the dataset is rather small for a definite conclusion to be drawn. Interestingly, the STD of background departures above 500 hPa are smaller despite the increase of wind speeds with height. The comparison of lidar and dropsonde winds (Weissmann *et al*., 2005b) did not show any correlation between wind speed and observation error either. For this reason the same error was assigned at all levels in contrast to the dropsonde error, which changes from 2 to 3 m s^{-1} with increasing height up to the tropopause.

The statistics show an improved representativeness of the observations by averaging the measurements towards the model resolution (Figures 2 and 3): on average, the STD of the lidar background departures is about 0.3 m s^{-1} smaller in the experiments with averaged data (Median, 4Rev) than in the experiment with a 10 km resolution (1Rev). Furthermore, the percentage of measurements that are rejected by the first-guess check and the VarQC decreases. However, there are no significant differences between the Median and 4Rev and at this stage it is not possible to determine which averaging method is better when using lidar measurements in NWP models.

In Median, 4Rev, and 4RStd the percentage of lidar measurements rejected by the first-guess check is smaller than the percentage of rejected aircraft measurements as a result of the lidar pre-processing. The percentage of measurements rejected by VarQC in contrast, was
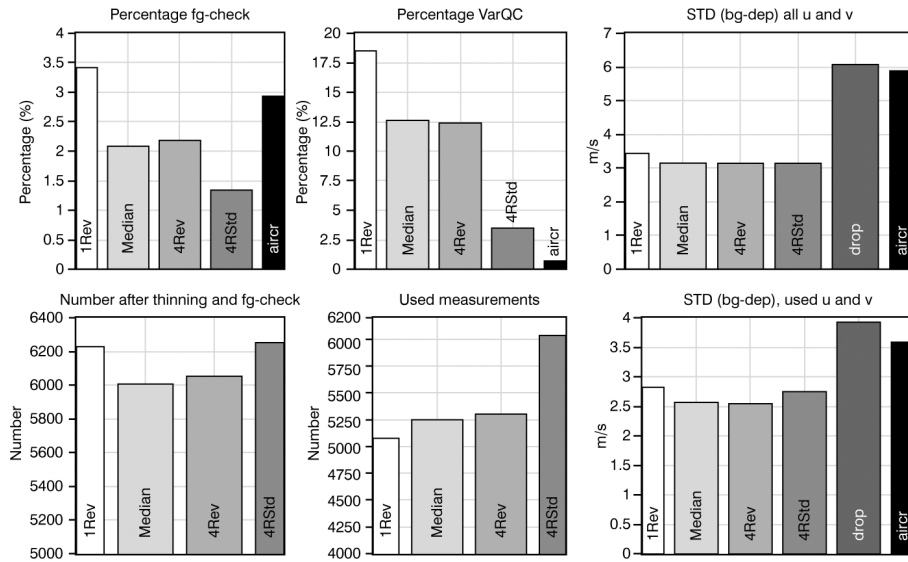
Figure 2. Assimilation statistics of different experiments: (top left) percentage of measurements discarded by the first-guess check; (top middle) percentage VarQC flagged; (top right) STD of the background departure of all measurements; (bottom left) number of all measurements after the first-guess check and the thinning procedure; (bottom middle) number of used measurements; (bottom right) STD of the background departure of all used measurements. The statistics for aircraft (aircr) and dropsonde (drop) measurements are shown for 4RStd and for the area 50 °N, 40 °W, 70 °N, 15 °W (main lidar operating area).
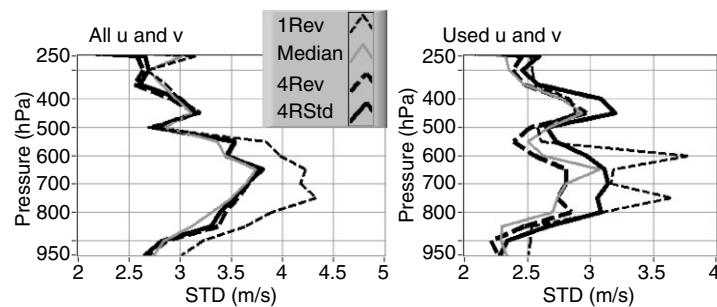
Figure 3. Vertical profile of the STD of lidar background departures for the four lidar experiments.

relatively high in all lidar experiments except the one with the increased observation error.

The data were thinned in the vertical from 100 m to the model level resolution. In Median, 4Rev, and 4RStd about 80% of the lidar data were thinned and the subsequent information in the vertical was lost. For future experiments a vertical averaging rather than thinning of the data is suggested.

4RStd has the largest number of used observations while 1Rev shows the smallest number of lidar data assimilated.

Figure 4 shows a vertical profile of the bias of lidar background departures. The largest bias is observed for the zonal wind component (u) at 300 hPa, presumably related to an underestimation of the jet level wind speed in the model (Tenenbaum, 1996; Cardinali *et al*., 2004b). At the other levels the bias was less than 1 m s^{-1}. The lidar measurements were shown to be almost unbiased by Weissmann *et al*. (2005b).

3.3.   Independent validation of lidar measurements

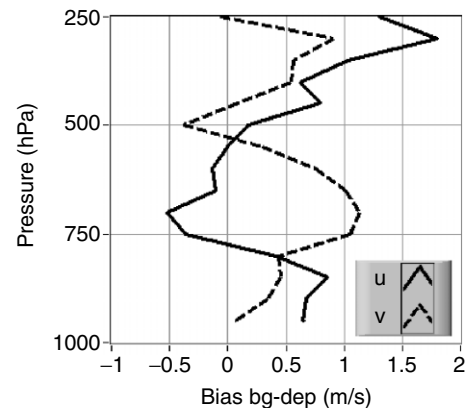In all lidar experiments, collocated dropsondes were used passively, which means the departures were computed

Figure 4. Vertical profile of the bias of background departures for lidar u, v wind components in 4RStd.

but not minimized (section 3.1). Thus, comparisons with dropsondes provide an independent validation of lidar measurements. The fact that the analysis departures for dropsondes are smaller than the background departures (Figure 5) means that lidar measurements lead to an analysis solution in the direction indicated by the dropsonde

observations. Furthermore, the analysis departure of drifting buoys surface pressure observations in the area of the lidar measurements was about 3% smaller in 4RStd than in the Control. The impact on the departures of other observations (aircraft, rawinsondes) was neutral, which is presumably owing to their distant locations.

3.4. Analysis impact

Observation influence in the analysis can be calculated during the minimization by finding the diagonal elements of the influence matrix (Hoaglin and Welsch, 1978; Cardinali *et al.*, 2004a). Because the observation error covariance matrix is diagonal, the diagonal elements of the influence matrix are bounded to 0 and 1. When the influence is 0, the analysis at the observation location is affected only by the background value (or pseudo-observation), while 1 means that only the observation counts for the final estimation at that location.

Figure 6 shows the vertically averaged influence of dropsonde and lidar observation on 22 November 2003. The influence of both observation types is high, because the measurements were taken in a data-sparse sensitive region. The mean influence of dropsonde and lidar wind observations is 0.45 and 0.63, respectively (Table II). In comparison, the mean observation influence of operational radiosondes in the northern hemisphere extratropics is ∼0.3, the mean influence of aircraft measurements and cloud-drift winds is 0.15 (Cardinali *et al.*, 2004a).

The mean lidar influence is 40% larger than the mean influence of dropsonde wind measurements and the information content of the lidar measurements expressed as degrees of freedom for signal (DFS; Wahba *et al.*, 1995) is nearly three times larger (Table II).
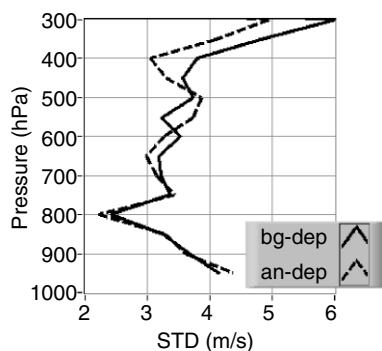

Figure 5. STD of the background (solid line) and analysis departure (dashed line) of dropsonde wind measurements in the experiment 4RStd.

Table II. Observation sensitivity for the analysis of 22 November 2003.

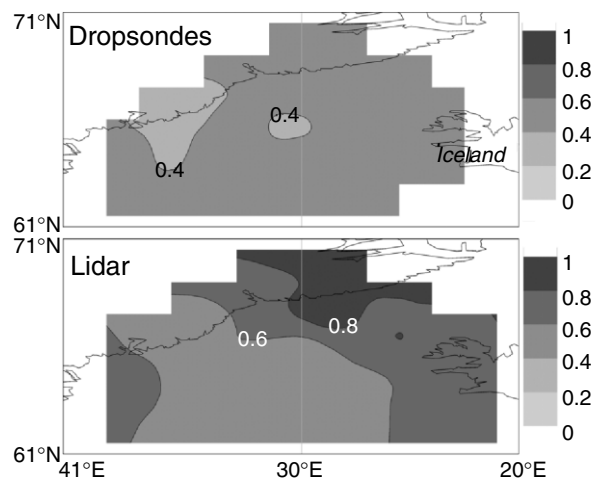|	Lidar u, v	Dropsonde u, v
Mean observation influence	0.63	0.45
Observation number (% of total)	758 (66)	388 (34)
Information content (% of total)	477.5 (73)	174.6 (27)

Figure 6. Vertically averaged observation influence on 22 November 2003 of (upper panel) dropsonde wind measurements from Drops and (lower panel) lidar wind measurements 4RStd. Numbers close to one indicate that the analysis is primarily based on the measurements, whereas numbers close to zero mean that the analysis is close to the background field.

3.5. Forecast impact

Figure 7 shows the difference, expressed as r.m.s. of forecast error, of 500 hPa geopotential height (Z) between 4RStd and Control for the period 15–28 November 2003. There is a clear reduction of the 2-day forecast error over the Atlantic Ocean and Northern Europe. At day 3 the reduction increases and propagates further to the east. The 4-day forecast error also decreases with the lidar measurements, and the main reduction of 4-day forecast errors is located over the Middle East, Northern and Eastern Europe. The improvement of 1-day forecasts (not shown) was fairly small and mostly restricted to the Atlantic Ocean.

To quantify the impact, the forecast error is averaged for Z, wind, and humidity for all levels over a box between 75°N, 15°W, 35°N, and 42.5°E (see Figure 7) and over 29 forecasts in the time period of the lidar deployment. All experiments with lidar data show a Z-error reduction for the 1–4 day forecast over Europe compared with the Control (Figure 8). The experiments Median and 4Rev do not have smaller forecast errors than 1Rev, although they show better departure statistics (section 3.2). The Z-error reduction increases with forecast time up to ∼2 m for the 4-day forecast. The increase is roughly proportional to the increase of forecast errors, and consequently, the relative reduction of Z-errors remains fairly steady around 3% for 2–4 day forecasts. On average, about 60% of the 29 forecasts improved (not shown). The spread between the different lidar experiments increases with time indicating the uncertainty due to the small sample of 29 forecasts. The impact after 5 days is not significant. A longer data sample would be necessary to investigate the long-term impact.

In many experiments the reduction of the forecast Z-error is largest in the upper troposphere (Figure 9). The maximum reduction of all experiments is 3 m at 300 hPa
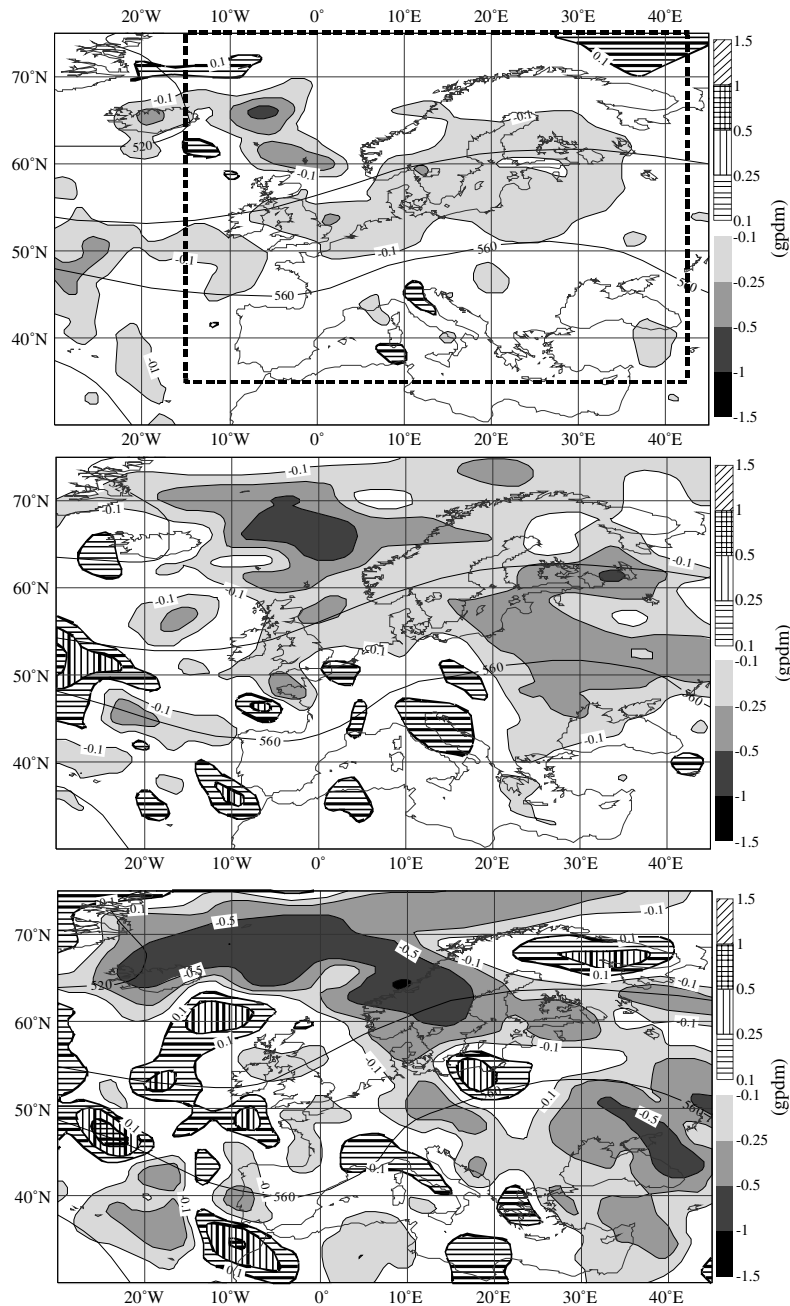
Figure 7. Difference of the r.m.s. Z-error at 500 hPa (gpdm) between 4RStd and Control of (top panel) 2-day forecasts, (middle) 3-day forecasts, and (bottom) 4-day forecasts between 15 and 28 November 2003 (28 forecasts). Negative values indicate a reduction of the forecast error. The dashed box in the upper panel shows the verification area 'Europe' used in this paper.

for the 4-day forecast. However, this increase with height is mainly due to the increase of the forecast error with height and the relative reduction does not correlate with height (not shown). The impact extended vertically up to 100 hPa by the background covariance matrix, which is well above the highest lidar measurements at ∼250 hPa. The maximum relative reduction of the 2–4 day forecast Z-error was up to 6% (Figures 8 and 9).

The experiment Drops also shows a reduction of the Z-error over Europe, mainly at day 1 and 2. At day 3 the reduction is fairly low and at day 4 some degradation above 700 hPa is observed. In general, experiments with lidar data show a larger and more sustained reduction of

the forecast error than Drops, which is consistent with their larger impact on the analysis.

Wind errors decrease in a similar way as the Z-errors (Figure 10). The relative reduction of wind errors in the troposphere is in the range of 0.4 to 3.3%, which is slightly smaller than the reduction of Z-errors. Furthermore, the lidar wind measurements improve the accuracy of the humidity forecasts (Figure 10) as an improved forecast of the wind and Z-field also leads to a more accurate forecast of humidity structures. The magnitude of the reduction is in the range of 0–2%. The reductions of 3–4 day forecast errors of humidity show maxima at 700 and 850 hPa, which are
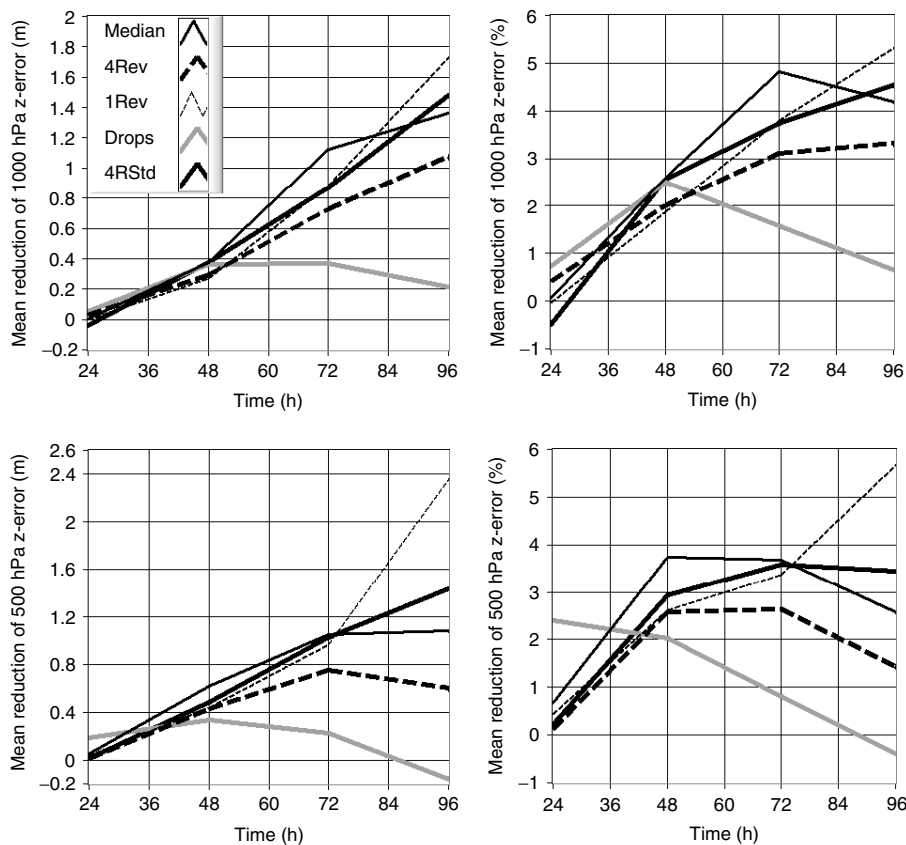
Figure 8. Mean reduction of Z-errors over Europe in different experiments compared with the Control run: (left panels) absolute reduction in metres; (right panels) relative reduction (in %) of the mean Z-error of the Control. Positive values correspond to a smaller error in the experiment than in the Control. The forecast error is averaged between all forecasts between 12 UTC 14 November and 12 UTC 28 November 2003 (29 forecasts). The verification area ('Europe') is shown in Figure 7.
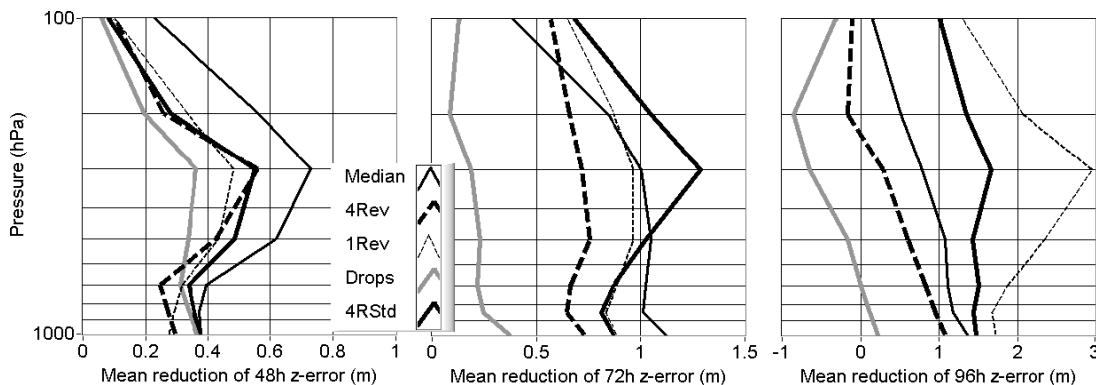

Figure 9. Vertical profile of the mean reduction of the Z-errors over Europe for forecast times of (left) 2 days, (middle) 3 days, and (right) 4 days. The mean reduction was calculated in the same way as in Figure 8 for all 29 forecasts between 14 and 28 November 2003.

presumably related to an improved prediction of frontal systems.

4. Conclusions and outlook

Despite recent advances in the use of satellite observations, there is a drastic shortage of wind measurements over the oceans (Tan and Andersson, 2005). This is a major deficiency in NWP as wind information has been shown to be of particular importance for representing

dynamical fields in the forecast (Cress and Wergen, 2001). Thus, Doppler lidar measurements with their high accuracy and resolution are a promising source of information to reduce errors in NWP models.

For the first time, airborne Doppler wind lidar observations have been assimilated in a global model in the framework of A-TReC. The present study evaluates the impact of these measurements. Four 2-week experiments were performed with different horizontal resolutions and assigned errors to optimize the assimilation of lidar data. The observations are averaged over a cycloid scan and are
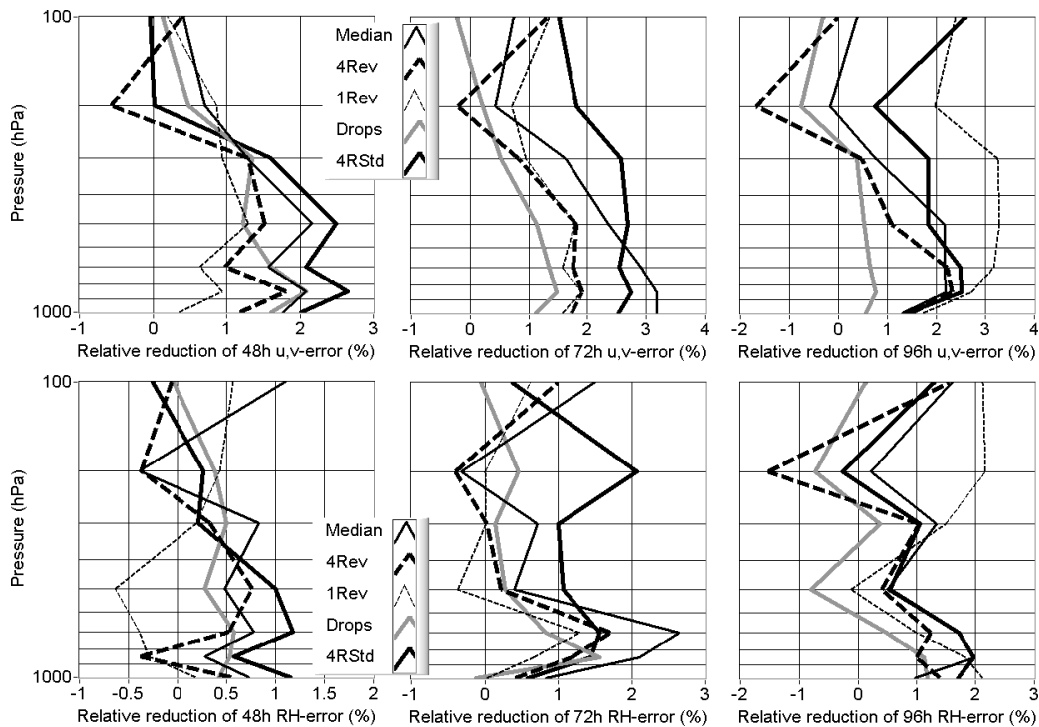
Figure 10. Mean relative reduction of forecast error for (top panels) u, v wind components and (bottom panels) the relative humidity over Europe. All values are as a percentage of the mean forecast error during the period 14–28 November 2003. (Left) 2-day forecasts; (middle) 3-day forecasts; and (right) 4-day forecasts. The relative reduction was calculated in the same way as in Figure 8, but for humidity and the wind components instead of Z.

therefore very representative of the wind field. The resulting assigned observation error variability (instrumental and representativeness error) is about half of most conventional observations. In consequence, the mean lidar observation influence (i.e. the impact on the analysis; Cardinali *et al.* 2004a) is determined to be 40% higher than the influence of dropsonde wind measurements. The lidar information content indicated by degrees of freedom for signal is shown to be three times larger than the dropsonde information content owing to the higher density (number) of lidar measurements.

In the performed experiments the 1–4 day forecast error of geopotential height, wind, and humidity over Europe reduces owing to the assimilation of the airborne Doppler lidar measurements over the Atlantic Ocean. The forecast of geopotential height improves at all levels from the surface up to 100 hPa. The mean reduction of the 2–4 day forecast error of geopotential height averaged over Europe and 29 forecasts from 14 to 28 November 2003 is around 3%. In all lidar experiments the improvement is consistent for different forecast variables (wind, humidity, and geopotential height) and forecast times.

Simmons and Hollingsworth (2002) documented that over the previous 11 years, the mean yearly reduction of the 3-day forecast error of the 500 hPa geopotential height was ~1 m for the Northern Hemisphere, which is comparable to the error reduction over Europe due to the assimilation of lidar data.

A recent data-denial study performed at ECMWF suggests that the theoretical upper limit of the average forecast error reduction over Europe reached with targeted observations is 2–5%. The impact of lidar data is therefore encouraging, even if it must be kept in mind that only 2 weeks of assimilation could be performed.

The results presented underline the importance of additional wind measurements above oceans and demonstrate the potential of Doppler lidars. These findings support the high expectations for the satellite-based Doppler lidar ADM-Aeolus, which the European Space Agency (ESA) plan to launch in 2008 (Stoffelen *et al.*, 2005). ADM-Aeolus will provide a global coverage (3000 profiles of LOS velocity a day), but the accuracy of ADM-Aeolus will only be half of the airborne lidar and the vertical resolution is 500–1000 m instead of 100 m. Given this difference, mounting Doppler lidars on commercial aircraft measuring high accuracy and high-resolution wind profiles over oceans should be investigated.

The experiment assimilating vertical profiles of wind and temperature from dropsondes for the same period shows only a mean reduction of ~1%, which is consistent with the smaller analysis impact of these observations compared with lidar data. Keeping in mind that the cost of lidar observations is comparable to the cost of dropsondes for the same flight time, the results given here show the potential of the airborne Doppler lidar for future observational campaigns

Specific investigation of the impact of lidar data in the verification regions established by the targeting techniques will be carried out in subsequent studies. Recently, experiments on the impact of differential absorption lidar (DIAL) water vapour measurements from different campaigns (Flentje *et al.*, 2005) have started, with promising

results. In the near future it will be also possible to operate the DIAL and Doppler lidar together on the DLR Falcon aircraft.

Acknowledgements

References

Andersson E, Järvinen H. 1999. Variational quality control. *Q. J. R. Meteorol. Soc.* **125**: 697–722.

Andersson E, Fisher M, Munro R, McNally A. 2000. Diagnosis of background errors for radiances and other observable quantities in a variational data assimilation scheme, and the explanation of a case of poor convergence. *Q. J. R. Meteorol. Soc.* **126**: 1455–1472.

Andersson E, Cardinali C, Isaksen L, Garcia-Mendez A. 2001. On the impact of frequent data in ECMWF's 4D-Var scheme: hourly surface pressure data, European profilers and profiling aircraft data. *Proceedings of the 8th ECMWF Workshop on Meteorological Operational Systems* 179–183.

Bastin S, Drobinski P, Dabas A, Delville P, Reitebuch O, Werner C. 2005. Impact of the Rhône and Durance Valleys on sea-breeze circulation in the Marseille area. *Atmos. Res.* **74**: 303–328.

Cardinali C, Isaksen L, Andersson E. 2003. Use and impact of automated aircraft data in a global 4DVAR data assimilation system. *Mon. Weather Rev.* **131**: 1865–1877.

Cardinali C, Pezzulli S, Andersson E. 2004a. Influence-matrix diagnostic of a data assimilation system. *Q. J. R. Meteorol. Soc.* **130**: 2767–2786.

Cardinali C, Rukhovet L, Tenenbaum J. 2004b. Jet stream analysis and forecast error using GADS observations in the DAO, ECMWF, and NCEP models. *Mon. Weather Rev.* **132**: 764–779.

Courtier P, Thépaut J-N, Hollingsworth A. 1994. A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.* **120**: 1367–1388.

Cress A, Wergen W. 2001. Impact of profile observations on the German Weather Service's NWP system. *Meteorol. Zeitschrift* **10**: 91–101.

Drobinski P, Bastin S, Guenard V, Caccia J-L, Dabas AM, Delville P, Protat A, Reitebuch O, Werner C. 2004. The Mistral at the exit of the Rhône Valley. *Q. J. R. Meteorol. Soc.* **130**: 1–21.

Flentje H, Dörnbrack A, Ehret G, Fix A, Kiemle C, Poberaj G, Wirth M. 2005. Water vapor heterogeneity related to tropopause folds over the North Atlantic revealed by airborne water vapor differential absorption lidar. *J. Geophys. Res.* **110**: D03115, doi: 10.1029/2004JD004957.

Grund CJ, Banta RM, George JL, Howell JN, Post MJ, Richter MA, Weickmann AM. 2001. High-resolution Doppler lidar for boundary layer and cloud research. *J. Atmos. Oceanic Technol.* **18**: 376–393.

Hoaglin DC, Welsch RE. 1978. The hatmatrix in regression and ANOVA. *American Statistician* **32**: 17–22.

Järvinen H, Undén P. 1997. Observation screening and first guess quality control in the ECMWF 3D-Var data assimilation system. *ECMWF Tech. Memo* 236.

Köpp F, Rahm S, Smalikho I. 2004. Characterization of aircraft wake vortices by 2-μm pulsed Doppler lidar. *J. Atmos. Oceanic Technol.* **21**: 194–206.

Langland R. 2005. Observation impact during the North Atlantic TReC-2003. *Mon. Weather Rev.* **133**: 2297–2309.

Rabier F. 2006. Review of results from ATReC. International Core Steering Committee for THORPEX, fifth session. 12 pages.

Rabier F, Clinker E, Courtier P, Hollingsworth A. 1996. Sensitivity of forecast errors of initial conditions. *Q. J. R. Meteorol. Soc.* **122**: 121–150.

Rabier F, Järvinen H, Klinker E, Mahfouf JF, Simmons A. 2000. The ECMWF operational implementation of four-dimensional variational assimilation. Part 1: Experimental results with simplified physics. *Q. J. R. Meteorol. Soc.* **126**: 1143–1170.

Reitebuch O, Volkert H, Werner C, Dabas A, Delville P, Drobinski P, Flamant PH, Richard E. 2003. Determination of air flow across the Alpine ridge by a combination of airborne Doppler lidar, routine radio-sounding, and numerical simulation. *Q. J. R. Meteorol. Soc.* **129**: 715–728.

Simmons AJ, Hollingsworth A. 2002. Some aspects of the improvement in skill of numerical weather prediction. *Q. J. R. Meteorol. Soc.* **128**: 647–677.

Smalikho I. 2003. Techniques of wind vector estimation from data measured with a scanning coherent Doppler lidar. *J. Atmos. Oceanic Technol.* **20**: 276–291.

Stoffelen A, Pailleux J, Källén E, Vaughan JM, Isaksen L, Flamant P, Wergen W, Andersson E, Schyberg H, Culoma A, Meynart R, Endemann M, Ingmann P. 2005. The Atmospheric Dynamics Mission for global wind field measurement. *Bull. Am. Meteorol. Soc.* **86**: 73–87.

Tan DGH, Andersson E. 2005. Simulation of the yield and accuracy of wind profile measurements from the Atmospheric Dynamics Mission (ADM-Aeolus). *Q. J. R. Meteorol. Soc.* **131**: 1737–1757.

Tenenbaum J. 1996. Jet stream winds: Comparisons of aircraft observations with analyses. *Weather and Forecasting* **11**: 188–197.

Veersé F, Thépaut J-N. 1998. Multiple-truncation incremental approach for four-dimensional variational data assimilation. *Q. J. R. Meteorol. Soc.* **124**: 1889–1908.

Wahba G, Johnso DR, Gao F, Gong J. 1995. Adaptive tuning of numerical weather prediction models: randomized GCV in three- and four-dimensional data assimilation. *Mon. Weather Rev.* **123**: 3358–3370.

Weissmann M, Braun FJ, Gantner L, Mayr G, Rahm S, Reitebuch O. 2005a. The Alpine mountain-plain circulation: airborne Doppler lidar measurements and numerical simulations. *Mon. Weather Rev.* **133**: 3095–3109.

Weissmann M, Busen R, Dörnbrack A, Rahm S, Reitebuch O. 2005b. Targeted observations with an airborne wind lidar. *J. Atmos. Oceanic. Technol.* **22**: 1706–1719.